

Robust Neighbor Discovery in Multi-Hop Multi-Channel Heterogeneous Wireless Networks*

Yanyan Zeng[†], K. Alex Mills, Shreyas Gokhale, Neeraj Mittal[‡], S. Venkatesan, and

R. Chandrasekaran

Department of Computer Science

The University of Texas at Dallas

Richardson, TX 75080, USA

E-mails: yanyan.bansal@gmail.com k.alex.mills@utdallas.edu

shreyas.gokhale@utdallas.edu neerajm@utdallas.edu venky@utdallas.edu

chandra@utdallas.edu

Abstract

An important first step when deploying a wireless ad hoc network is *neighbor discovery* in which every node attempts to determine the set of nodes it can communicate within one wireless hop. In the recent years, *cognitive radio (CR)* technology has gained attention as an attractive approach to alleviate spectrum congestion. A cognitive radio transceiver can operate over a wide range of frequencies possibly spanning multiple frequency bands. A cognitive radio node can opportunistically utilize unused wireless spectrum without interference from other wireless devices in its vicinity. Due to spatial variations in frequency usage and hardware variations in radio transceivers, different nodes in the network may perceive different subsets of frequencies available to them for communication. This *heterogeneity* in the available channel sets across the network increases the complexity of solving the neighbor discovery problem in a cognitive radio network. In this work, we design and analyze several randomized algorithms for neighbor

*A preliminary version of the work has appeared earlier in the Proceedings of the 31st International Conference on Distributed Computing Systems (ICDCS), 2011 [45].

[†]Yanyan Zeng is currently working at JPMorgan Chase.

[‡]This work was supported, in part, by the National Science Foundation (NSF) under grant number CNS-1115733.

discovery in such a (heterogeneous) network under a variety of assumptions (*e.g.*, maximum node degree known or unknown) for both synchronous and asynchronous systems under minimal knowledge. We also show that our randomized algorithms are naturally suited to tolerate unreliable channels and adversarial attacks.

Key words: multi-hop multi-channel wireless network, cognitive radio technology, heterogeneous channel availability, neighbor discovery, randomized algorithm, asynchronous system, clock drift, lossy channel, jamming attack

1 Introduction

Neighbor discovery is an important step in forming a *self-organizing* wireless ad hoc network without any support from an existing communication infrastructure [8,9,19,20,29,42,51,54,57,58,63]. When deployed, nodes initially have no prior knowledge of other nodes that they can communicate with directly.

The results of neighbor discovery can be used to solve other important communication problems such as medium access control [7, 22], routing [15], broadcasting [50, 54], clustering [16, 32, 40], collision-free scheduling [26, 27], spanning tree construction [21], and topology control [37, 62]. Many algorithms for solving these problems implicitly assume that all nodes know their one-hop and sometimes even two-hop neighbors. Many location-based routing protocols (*e.g.* localized routing in Vehicular Ad hoc Networks (VANETs)) use the position of neighboring nodes to make routing/forwarding decisions. The neighborhood information is also used to update the reachability status of nodes. A better neighbor discovery algorithm, which uses fewer messages and has higher accuracy, can be used to improve the performance of location-based routing algorithms [10–12]. Neighborhood information helps reduce the cost of flooding in multicast tree construction using flooding algorithms [39]. Mobile sensing applications, ranging from mobile social networking to proximity-based gaming, involve collection and sharing of data among nearby users. The success of these applications depends on neighborhood information [20]. Neighbor discovery is extremely important in Underwater Acoustic (UWA) Networks and needs to be done frequently because nodes may move proactively due to the unpredictable underwater currents [68]. More details of how the results of neighbor discovery can be used to solve other communication problems can be found

in [10, 11].

Cognitive radio (CR) technology has emerged as a promising approach for improving spectrum utilization efficiency and meeting the increased demand for wireless communications [13]. A CR node can scan a part of the wireless spectrum, and identify unused or underutilized channels in the spectrum [1, 65]. CR nodes in a network can then use these channels opportunistically for communication among themselves even if the channels belong to licensed users. The licensed users are referred to as the *primary users*, and CR nodes are referred to as the *secondary users*. (Of course, when a primary user arrives and starts using its channel, the secondary users have to vacate the channel.) Due to spatial variations in frequency usage, hardware variations in radio transceivers and uneven propagation of wireless signals, different nodes in the network may perceive different subsets of frequencies available to them for communication. This gives rise to a *multi-hop, multi-channel, heterogeneous wireless network*, abbreviated as *M²HeW network*. The focus of this work is on solving the neighbor discovery problem in an M²HeW network.

A large number of neighbor discovery algorithms have been proposed in the literature. Most of the algorithms suffer from one or more of the following limitations: (i) all nodes are assumed to be synchronized (synchronous system), (ii) the entire network is assumed to operate on a single channel (single-channel network), (iii) all nodes are assumed to be able to communicate with each other (single-hop network), (iv) all channels are assumed to be available to all nodes (homogeneous network), or (v) the algorithm is only evaluated experimentally (no theoretical guarantees provided). A more detailed discussion of the related work can be found in Section 7.

Our Contributions: Our main contribution in this work is *two* randomized neighbor discovery algorithms for an *M²HeW network* when the system is *asynchronous* that guarantee success with arbitrarily high probability. The first algorithm assumes that nodes know a good upper bound on the maximum degree of any node in the network. The second algorithm does not make any such assumption. Both algorithms only assume that the maximum drift rate of the clock of any node is bounded, with the second algorithm assuming a tighter but unknown bound whose value depends on various system parameters. None of the algorithms require clocks of different nodes to be synchronized. In fact, clocks of any two nodes may have arbitrary skew with respect to each other. Other advantages of our algorithms are as follows: (i) nodes do not need to agree

on a universal channel set, and (ii) the running time of an algorithm depends on the “degree of heterogeneity” in the network; the running time decreases as the available channel sets become more homogeneous.

Our algorithms for an asynchronous system are based on those for a synchronous system. Therefore, as additional contributions, we also present a suite of randomized neighbor discovery algorithms for an M^2HeW network when the system is *synchronous* under a variety of assumptions such as: (i) whether all the nodes start executing the neighbor discovery algorithm at the same time or not, and (ii) whether nodes are aware of an estimate on the upper bound on the maximum degree of any node in the network or not.

We believe that our approach for transforming a *state-less* algorithm developed for a synchronous M^2HeW network to work for an asynchronous M^2HeW network can also be applied to other important communication problems in an M^2HeW network with running time increasing by only a constant factor.

We show that our randomized algorithms can easily tolerate unreliable channels. We also prove that, our algorithms, with minor modification in the asynchronous case, are tolerant to jamming attacks by a reactive but “memory-less” jammer under certain assumption. The running time of our algorithms, when subject to a jamming attack, increases by at most a *constant* factor. In fact, for sufficiently large values of system parameters (namely, number of nodes and number of channels), the running time increases only by a factor of at most two in the worst-case.

Organization: The rest of the manuscript is organized as follows. We describe our system model for a multi-hop multi-channel heterogeneous wireless network in Section 2. For ease of exposition, we first present a suite of randomized neighbor discovery algorithms for a *synchronous* system under a variety of assumptions and analyze their time complexity in Section 3. We then present two randomized neighbor discovery algorithms for an *asynchronous* system, which are derived from their synchronous counterparts, and analyze their complexity in Section 4. We discuss several extensions to our algorithms to enhance their applicability and improve their robustness in Section 6. Finally, in Section 7, we give a comparison of our contributions with existing research and also examine other related work done on neighbor discovery.

2 System Model

We assume a multi-hop multi-channel heterogeneous wireless (M²HeW) network consisting of one or more radio nodes. Let N denote the total number of radio nodes. Nodes do not know N . Each radio node is equipped with a transceiver (transmitter-receiver pair), which is capable of operating over multiple frequencies or channels. However, at any given time, a transceiver can operate (either transmit or receive) over a single channel only. Further, a transceiver cannot transmit and receive at the same time. Transceivers of different nodes need not be identical; the set of channels over which a transceiver can operate may be different for different transceivers.

Different nodes in a network may have different sets of channels available for communication. For example, in a cognitive radio network (a type of M²HeW network), each node can scan the frequency spectrum and identify the subset of unused or under-used portions of the spectrum even those that have been licensed to other users or organizations [13]. A node can potentially use such frequencies to communicate with its neighbors until they are reclaimed by their licensed (primary) users [13]. Due to spatial variations in frequency usage/interference and hardware variations in radio transceivers, different nodes in the network may perceive different subsets of frequencies available to them for communication. We refer to the subset of frequencies or channels that a node can use to communicate with its neighbors as the *available channel set* of the node. For a node u , we use $\mathcal{A}(u)$ to denote its available channel set. We use S to denote the size of the largest available channel set, that is, $S = \max_u |\mathcal{A}(u)|$. Note that nodes do not know S .

We say that a node v is a *neighbor* of node u on a channel c if u can reliably receive any message transmitted by v on c provided no other node in the network is transmitting on c at the same time, and vice versa. We assume that the communication graph is *symmetric* because unidirectional neighborhood relationships are expensive and impractical to use in wireless networks [47] (although our algorithms work for asymmetric communication graphs as well). For a node u and a channel $c \in \mathcal{A}(u)$, we use $\Delta(u, c)$ to denote the number of neighbors, also known as *degree*, of u on c . We use Δ denote the maximum degree of any node on any channel, that is, $\Delta = \max_u \max_{c \in \mathcal{A}(u)} \Delta(u, c)$.

Note that, if nodes u and v are neighbors of each other on some channel, then u has to discover v and v has to discover u separately. It is convenient to assume two separate links—one from u to v and another from v to u . We use (u, v) to denote the link from u to v . We refer to the set of

channels on which the link (u, v) can operate as the *span* of (u, v) and denote it by $\text{span}(u, v)$. Note that $\text{span}(u, v) \subseteq \mathcal{A}(u) \cap \mathcal{A}(v)$. We refer to the ratio of $|\text{span}(u, v)|$ to $|\mathcal{A}(v)|$ as the *span-ratio* of the link (u, v) . Note that the span-ratio of any link lies between $\frac{1}{5}$ and 1. Further, the span-ratio of (u, v) may be different from the span-ratio of (v, u) (because $|\mathcal{A}(u)|$ may be different from $|\mathcal{A}(v)|$). We use ρ to denote the minimum span-ratio among all links. Note that nodes do not know ρ . Intuitively, ρ can be viewed as a measure of the *degree of homogeneity* in the network—the larger the value, the more homogeneous the network is in terms of channel availability for nodes and links. The running time of our algorithms is *inversely* proportional to ρ . When all links can operate on all available channels (an assumption made frequently in the literature), $\rho = 1$, which minimizes the running time of our algorithms. In general, the more heterogeneous the network is, the larger is the running time of our algorithms.

If nodes v and w are neighbors of node u on channel c and both v and w transmit on c at the same time, then their transmissions *collide* at u . If u is listening on c when v and w are transmitting on c at the same time, then u hears only noise. We do not assume that nodes can detect collisions, that is, distinguish between background noise and collision noise.

The set of channels over which an M²HeW network can operate may have very different propagation characteristics. For example, a cognitive radio transceiver may be capable of operating over a wide range of frequencies (*e.g.*, as low as 100 MHz to as high as 2.5 GHz [14]). Radio waves at different frequencies undergo different angles of refraction and therefore propagate in different ways. As a result, two radio nodes u and v that are neighbors on some channel c may not be neighbors on another channel d even if both u and v have d in their available channel sets. This may happen, for example, if c is a channel in the 900 MHz band and d is a channel in the 2.4 GHz band; a lower frequency has stronger capability to penetrate through different materials and therefore has a longer transmission range than a higher frequency. For ease of exposition, we first describe our algorithms assuming that all frequencies or channels have identical propagation characteristics. This implies that if a communication link from node u to node v can operate over some channel $c \in \mathcal{A}(u) \cap \mathcal{A}(v)$ then it can also operate over all channels in $\mathcal{A}(u) \cap \mathcal{A}(v)$. We later discuss how to extend our algorithms to handle diverse propagation characteristics of different frequencies.

We use “log” to refer to the logarithm to the base 2 and “ln” to refer to the natural logarithm. In this paper, we investigate the neighbor discovery problem both when the system is synchronous

and when the system is asynchronous. We now describe how we model each type of system.

Synchronous System: We assume that the execution of the system is divided into *synchronized time-slots*. In each time slot, each node can be in one of the following three modes: (i) *transmit mode* on some channel in its available channel set, (ii) *receive mode* on some channel in its available channel set, or (iii) *quiet mode* when the transceiver is shut-off. Nodes may or may not start at the same time.

Asynchronous System: We assume that every node is equipped with a clock. Clocks of different nodes are not required to be synchronized. For a clock C and time t , we use $C(t)$ to denote the value of C at t . For a node u , we use C_u to denote the clock of u . The clock of a node may have non-zero drift and the drift rate may change over time. The drift rate of a clock C at time t is given by $\frac{dC}{dt} - 1$. If C is an ideal clock, then, $\forall t, \Delta t \geq 0$, $C(t + \Delta t) - C(t) = \Delta t$; thus $\frac{dC}{dt} = 1$. At any given time, different clocks may have different drift rates. Further, the drift rate of the same clock may change over time both in magnitude and sign. We do, however, assume that the magnitude of the maximum drift rate is *bounded* and let the bound be denoted by δ . This implies that $\forall t, \Delta t \geq 0$,

$$(1 - \delta)\Delta t \leq C(t + \Delta t) - C(t) \leq (1 + \delta)\Delta t \quad (1)$$

For an ideal clock, $\delta = 0$. In practice, δ is quite small (*e.g.*, 10^{-6} seconds/second for an ordinary quartz clock).

3 Neighbor Discovery in Synchronous Systems

We first describe our algorithms for neighbor discovery assuming that all nodes start executing the algorithm at the same time. We then relax this assumption and describe our algorithms for neighbor discovery when nodes may start executing the algorithm at different times. Later, in Section 4, we use the algorithms for variable start times to derive randomized neighbor discovery algorithms for an asynchronous system.

```

// Algorithm for node  $u$ 
Input: An upper bound on maximum node degree, say  $\Delta_{est}$ .
Output: The set of neighbors along with the subset of channels that are common with the neighbor.
1 while true do
  // execute a stage
2   for  $i \leftarrow 1$  to  $\lceil \log(\Delta_{est}) \rceil$  do
    // time-slot  $i$  of the current stage
    // select a channel
3     $c \leftarrow$  channel selected uniformly at random from  $\mathcal{A}(u)$ ;
    // compute the transmission probability
4     $p \leftarrow \min\left(\frac{1}{2}, \frac{|\mathcal{A}(u)|}{2^i}\right)$ ;
5    tune the transceiver to  $c$ ;
6    switch to transmit mode with probability  $p$  and receive mode with probability  $1 - p$ ;
7    if (in transmit mode) then
8      | transmit a message containing  $\mathcal{A}(u)$ ;
9    else if (heard a clear transmission) then
10   | let the received message be sent by node  $v$  containing set  $A$ ;
11   | add  $\langle v, A \cap \mathcal{A}(u) \rangle$  to the set of neighbors;
12   end if
13   end for
14 end while

```

Algorithm 1: Neighbor discovery algorithm for a synchronous system with identical start times and knowledge of upper bound on maximum node degree.

3.1 Identical Start Times

We first assume that the nodes know some upper bound on maximum node degree and present a neighbor discovery algorithm. The bound need not be tight and, in fact, may be quite loose (since the dependence is logarithmic on the value of the upper bound.) But all nodes should agree on a *common* upper bound. We then relax this restriction and describe an algorithm for the case when such knowledge is not available.

3.1.1 Knowledge of Loose Upper Bound on Maximum Node Degree

Let Δ_{est} denote an upper bound on the maximum node degree as known to all nodes. The execution of the algorithm is divided into *stages*. Each stage consists of $\lceil \log(\Delta_{est}) \rceil$ time-slots. In each time-slot of a stage, a node randomly chooses a channel from its available channel set and transmits on that channel with a certain probability (and listens on that channel with the remaining probability). Specifically, in time-slot i of a stage, where $1 \leq i \leq \lceil \log(\Delta_{est}) \rceil$, a node u transmits on the selected channel, say c , with probability $\min\left(\frac{1}{2}, \frac{|\mathcal{A}(u)|}{2^i}\right)$ and listens on c with the remaining probability.

A formal description of the algorithm is shown in Algorithm 1. We next analyze its running time. Consider a node u and let node v be its neighbor on a channel c . Observe that $1 \leq \Delta(u, c) \leq$

$\Delta \leq \Delta_{est}$. Let $k = \max(1, \lceil \log \Delta(u, c) \rceil)$. Clearly, we have:

$$2^{k-1} \leq \Delta(u, c) \leq 2^k \quad (2)$$

We say that a time-slot t *covers* the link (v, u) on channel c if during t : (i) v transmits on c , (ii) u listens on c , and (iii) no other neighbor of u transmits on c . The three conditions collectively ensure that u receives a clear message from v (on c during t). We say that a stage s covers the link (v, u) on channel c if some time-slot of s covers (v, u) on c ; also, s covers (v, u) if s covers (v, u) on some $c \in \text{span}(v, u)$. Finally, we say that a sequence of stages covers the link (v, u) if some stage of the sequence covers (v, u) .

Consider a stage s and let τ denote the time-slot of s that satisfies (2). We first compute the probability that (v, u) is covered on c during τ . Let $A(\tau, c)$, $B(\tau, c)$, $C(\tau, c)$ denote the three events corresponding to the three conditions (i), (ii) and (iii), respectively, required for coverage. Note that the three events are *mutually independent*. Therefore we first compute the probability of occurrence of events $A(\tau, c)$, $B(\tau, c)$ and $C(\tau, c)$ separately. We then compute the probability that s covers (v, u) . The steps involved in computing the probabilities are fairly standard and have been moved to the appendix. We only state the results here. We have:

$$\Pr\{A(\tau, c)\} \geq \frac{1}{2 \max(S, \Delta)} \quad (3)$$

$$\Pr\{B(\tau, c)\} \geq \frac{1}{2|\mathcal{A}(u)|} \quad (4)$$

$$\Pr\{C(\tau, c)\} \geq \frac{1}{4} \quad (5)$$

$$\Pr(s \text{ covers } (v, u)) \geq \frac{\rho}{16 \max(S, \Delta)} \quad (6)$$

Now, consider a sequence of $M = \frac{16 \max(S, \Delta)}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$ stages, where ϵ is the probability that neighbor discovery does not complete successfully. We show that the probability that the link (v, u) is not covered within M stages is at most $\frac{\epsilon}{N^2}$. The steps for computing the probability are fairly standard and have been moved to the appendix. Formally,

$$\Pr((v, u) \text{ is not covered within } M \text{ stages}) \leq \frac{\epsilon}{N^2} \quad (7)$$

Finally, we have:

$$\begin{aligned}
& \Pr(\text{neighbor discovery does not finish within } M \text{ stages}) \\
&= \Pr(\text{some link is not covered within } M \text{ stages}) \\
&\leq (\text{number of links in the network}) \times \frac{\epsilon}{N^2} \leq \epsilon
\end{aligned}$$

Therefore, we have the following theorem:

Theorem 1. *Algorithm 1 guarantees that each node discovers all its neighbors on all channels within $O\left(\frac{\max(S, \Delta)}{\rho} \log(\Delta_{est}) \log\left(\frac{N}{\epsilon}\right)\right)$ time-slots with probability at least $1 - \epsilon$.*

3.1.2 No Knowledge of Maximum Node Degree

One way to derive a neighbor discovery algorithm when knowledge about maximum node degree is not available is as follows. Starting with an estimate of one for the maximum node degree, repeatedly run an instance of the algorithm that assumes knowledge about the maximum degree for a certain number of time-slots with geometrically increasing values for the estimate. This approach, for example, is used by Vasudevan *et al.* in [58] to eliminate the assumption about network size knowledge made by some of their neighbor discovery algorithms. The approach has the advantage that the time complexity of the knowledge-unaware algorithm is asymptotically same as that of the knowledge-aware algorithm. This approach cannot be used here because it requires computing the exact number of time-slots for which an instance of the knowledge-aware algorithm ought to be run such that, in case the estimate is correct, the neighbor discovery completes with a desired success probability. Computing the number of time-slots for our (knowledge-aware) algorithm requires nodes to *a priori* know the values for other system parameters, namely N , S and ρ , whose values may not be known in advance.

We instead employ the following approach used in [46]. Starting with an estimate of two, we repeatedly execute an instance of stage with sequentially increasing values for the estimate.

A formal description of the algorithm is shown in Algorithm 2. It can be easily verified that, once d becomes at least Δ , each stage thereafter contains a time-slot that satisfies (2). To reach the success probability of at least $1 - \epsilon$, from the analysis of Algorithm 1, it suffices for an execution to

```

// Algorithm for node  $u$ 
Output: The set of neighbors along with their available channel set
1  $d \leftarrow 2$ ;
2 while  $true$  do
3   | execute an instance of stage described in Algorithm 1 with  $\Delta_{est}$  set to  $d$ ;
4   |  $d \leftarrow d + 1$ ;
5 end while

```

Algorithm 2: Neighbor discovery algorithm for a synchronous system with identical start times and no knowledge of maximum node degree.

contain $M = \frac{16 \max(S, \Delta)}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$ stages each of which consists of a time-slot that satisfies (2). In other words, neighbor discovery completes successfully with probability at least $1 - \epsilon$ within $\Delta + M$ stages. Note that $M = \Omega(\Delta)$. Therefore we have:

Theorem 2. *Algorithm 2 guarantees that each node discovers all its neighbors on all channels within $O(M \log M)$ time-slots with probability at least $1 - \epsilon$, where $M = \frac{16 \max(S, \Delta)}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$.*

3.2 Variable Start Times

We first assume that the nodes know some upper bound on maximum node degree and present a neighbor discovery algorithm. We then relax this restriction and describe an algorithm for the case when such a knowledge is not available. Let T_s denote the time by which all nodes have initiated neighbor discovery.

3.2.1 Knowledge of Good Upper Bound on Maximum Node Degree

We assume that nodes know a “good” upper bound on the maximum node degree. Although the algorithm works even if the upper bound is loose, the running time of the algorithm may be too large since it depends linearly on the value of the upper bound.

The main idea behind our algorithm is to ensure that the transmission probability of a node is the *same* for every time-slot (but may be different for different nodes). This allows us to prove that a given link is covered in a time slot with “sufficiently high” probability. Let Δ_{est} denote an upper bound on the maximum node degree as known to all nodes. In each time-slot, a node u randomly selects a channel from its available channel set, say c . It then transmits on c with probability $\min \left(\frac{1}{2}, \frac{|A(u)|}{\Delta_{est}} \right)$ and listens on c with the remaining probability. A formal description of the algorithm is shown in Algorithm 3.

```

// Algorithm for node u
Input: An upper bound on maximum node degree, say  $\Delta_{est}$ .
Output: The set of neighbors along with the subset of channels that are common with the neighbor.

// compute the transmission probability; same for all time-slots
1  $p \leftarrow \min\left(\frac{1}{2}, \frac{|\mathcal{A}(u)|}{\Delta_{est}}\right)$ ;
2 while true do
    // select a channel
3      $c \leftarrow$  channel selected uniformly at random from  $\mathcal{A}(u)$ ;
4     tune the transceiver to  $c$ ;
5     switch to transmit mode with probability  $p$  and receive mode with probability  $1 - p$ ;
6     if (in transmit mode) then
7         | transmit a message containing  $\mathcal{A}(u)$ ;
8     else if (heard a clear transmission) then
9         | let the received message be sent by node  $v$  containing set  $A$ ;
10        | add  $\langle v, A \cap \mathcal{A}(u) \rangle$  to the set of neighbors;
11    end if
12 end while

```

Algorithm 3: Neighbor discovery algorithm for a synchronous system with variable start times and knowledge of upper bound on maximum node degree.

For the analysis of the running time, as in the case of Algorithm 1, we can compute the probability of occurrence of events $A(\tau, c)$, $B(\tau, c)$ and $C(\tau, c)$. We have:

$$\begin{aligned}
 \Pr\{A(\tau, c)\} &= \frac{1}{|\mathcal{A}(v)|} \times \min\left(\frac{1}{2}, \frac{|\mathcal{A}(v)|}{\Delta_{est}}\right) \\
 &= \frac{1}{\max\{2|\mathcal{A}(v)|, \Delta_{est}\}} \\
 &\geq \frac{1}{\max(2S, \Delta_{est})}
 \end{aligned} \tag{8}$$

It can be verified that the inequalities for $\Pr\{B(\tau, c)\}$ in (4) and $\Pr\{C(\tau, c)\}$ in (5) are still valid (although the proofs have to be slightly modified). Using an analysis similar to that for Algorithm 1, we can prove the following result:

Theorem 3. *Algorithm 3 guarantees that each node discovers all its neighbors on all channels within $O\left(\frac{\max(2S, \Delta_{est})}{\rho} \log\left(\frac{N}{\epsilon}\right)\right)$ time-slots of T_s with probability at least $1 - \epsilon$.*

Note that we no longer have a factor of $O(\log(\Delta_{est}))$ in the time complexity because we do not have stages any more.

```

// Algorithm for node  $u$ 
Output: The set of neighbors along with the subset of channels that are common with the neighbor.
1  $i \leftarrow 1$ ;
2 while true do
    // epoch  $i$ 
3   for  $j \leftarrow 1$  to  $i + 1$  do
    // phase  $j$  in epoch  $i$ 
4     execute an instance of Algorithm 3 with  $\Delta_{est} = 2^j$  for  $2^i$  time-slots;
5   end for
6    $i \leftarrow i + 1$ ;
7 end while

```

Algorithm 4: Neighbor discovery algorithm for a synchronous system with variable start times (no knowledge assumed).

3.2.2 No Knowledge of Maximum Node Degree

We refer to the difference in the start times of two nodes as *slack*, and use Θ to denote the maximum slack between any two nodes (expressed in terms of the number of time-slots). Also, note that no node knows Θ . In fact, we present a neighbor discovery algorithm in which no node knows any of the other system parameters, namely, S , Δ , ρ , N and ϵ . And, yet the algorithm completes neighbor discovery within a “short” period with high probability.

The execution of the algorithm is divided into *epochs*. Each epoch is further divided into *phases*. Each phase consists of a certain number of time-slots. Specifically, epoch i , for $i = 1, 2, 3, \dots$, consists of $i + 1$ phases. In phase j of epoch i , where $1 \leq j \leq i + 1$, we execute an instance of Algorithm 3 with Δ_{est} set to 2^j for a *fixed* number of time-slots given by 2^i . Note that the number of phases in an epoch increases linearly but the length of a phase in an epoch increases geometrically with the epoch number. Also, note that all phases in an epoch are of the same length but use geometrically increasing values for estimate of the upper bound on maximum node degree. A formal description of the algorithm is shown in Algorithm 4.

In order to show that Algorithm 4 guarantees that neighbor discovery completes with high probability within a certain number of time-slots, it is sufficient to identify a phase in an epoch such that: (i) nodes use a *sufficiently large value as an estimate* for maximum node degree during this phase, and (ii) this phase for all nodes *overlaps for a sufficiently long time* so as to allow all links to be discovered.

Let Δ_0 denote the smallest power of two greater than or equal to δ . Further, let $M =$

$\frac{16 \max(S, \Delta_0)}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$. Given a node u , we refer to the phase numbered $\log \Delta_0$ of the epoch numbered $\lceil \log(M + \Theta) \rceil$ as the *potent phase* of node u . (We call it potent because we show that neighbor discovery is guaranteed to complete with high probability in this phase.) In its potent phase, each node executes an instance of Algorithm 3 with Δ_{est} set to Δ_0 for at least $M + \Theta$ time-slots. Let T_b denote the *latest time* at which some node *begins* executing its potent phase. Likewise, let T_e denote the *earliest time* at which some node *ends* executing its potent phase. Note that the time difference between when two nodes *start the same phase* is same as the slack between the two nodes (the time difference between when two nodes start the neighbor discovery algorithm). Since the maximum slack between any two nodes is Θ , clearly, $T_e - T_b \geq M$. This implies that, during the time period from T_b to T_e , the following holds: (i) the potent phases of all nodes *mutually overlap* in time, and (ii) the time-period contains *at least M time-slots* from the potent phase of every node. We refer to the period of time from T_b to T_e as the *potent period*. During the potent period, each node uses the same estimate for the upper bound on maximum node degree, namely Δ_0 . Using an analysis similar to that for the previous algorithm, it can be shown that a single time-slot during the potent period covers a given link with probability at least $\frac{\rho}{16 \max(S, \Delta_0)}$. This in turn implies that the probability that neighbor discovery completes successfully during the potent period is at least $1 - \epsilon$.

Note that epoch i contains $(i + 1) \cdot 2^i$ time-slots. Thus, the total number of time-slots contained in epochs 1 through i is given by $i \cdot 2^{i+1}$ (proof is given in the appendix). As a result, we have:

Theorem 4. *Algorithm 4 guarantees that each node discovers all its neighbors on all channels within $O((M + \Theta) \log(M + \Theta))$ time-slots of T_s with probability at least $1 - \epsilon$, where $M = \frac{16 \max(S, \Delta_0)}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$.*

4 Neighbor Discovery in Asynchronous System

We first describe an algorithm that assumes that nodes know a good upper bound on maximum node degree. We then describe a more general algorithm in which no such knowledge is assumed. The second algorithm, however, requires a tighter bound on the clock drift rate than the first algorithm. Both algorithms are derived from their synchronous counterparts in which different nodes can start at different times.

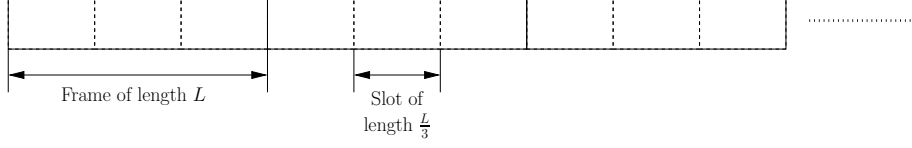


Figure 1: Execution of a node with respect to its local clock. All frames are of length L . All slots are of length $\frac{L}{3}$.

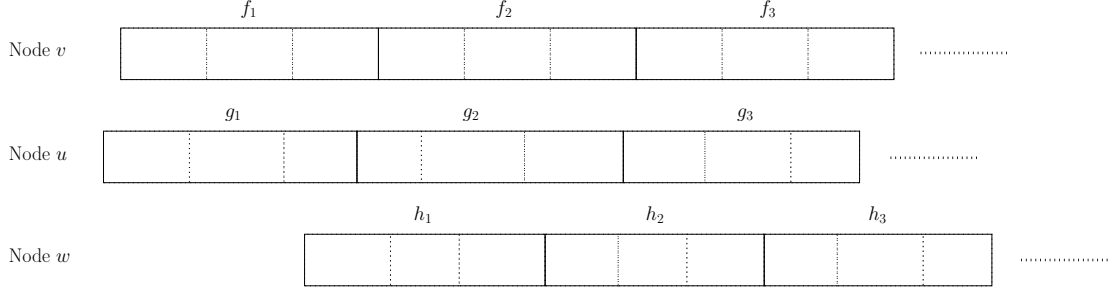


Figure 2: Execution of the network with respect to real time. Frames may be of different lengths (even within the same node). Slots may be of different lengths (even within the same frame).

4.1 Knowledge of Good Upper Bound on Maximum Node Degree

Our neighbor discovery algorithm for an asynchronous system is based on our neighbor discovery algorithm for a synchronous system with variable start times (Algorithm 3). Let Δ_{est} denote an upper bound on the maximum node degree as known to all nodes. In addition, our algorithm makes the following assumption about the maximum drift rate of the clock of any node δ :

Assumption 1. *The maximum drift rate of the clock of any node is bounded by $\frac{1}{7}$ seconds/second.*

The offset or skew between clocks of any two nodes may be arbitrarily large and, in fact, may grow with time. We now describe how to extend Algorithm 3 to solve the neighbor discovery algorithm in an asynchronous system.

Each node divides its time into equal-sized *frames*. Frames of different nodes are *not* required to be synchronized and may, in fact, be *misaligned*. A node measures the duration of a frame using its *local* clock. The length of a frame *as measured by a node using its local clock* is same for all nodes, say L . Note that, because of the clock drift, duration of a frame, when projected on real time, may be different from L —shorter than L in case of positive drift and longer than L in case of negative drift. Specifically, it can be verified that the length of a frame in real time lies in the

```

// Algorithm for node  $u$ 
Input: An upper bound on maximum node degree, say  $\Delta_{est}$ .
Output: The set of neighbors along with the subset of channels that are common with the neighbor.

// compute the transmission probability
1  $p \leftarrow \min\left(\frac{1}{2}, \frac{|\mathcal{A}(u)|}{3\Delta_{est}}\right)$ ;
2 while true do
    // select a channel
3      $c \leftarrow$  channel selected uniformly at random from  $\mathcal{A}(u)$ ;
4     tune the transceiver to  $c$ ;
5     switch to transmit mode with probability  $p$  and receive mode with probability  $1 - p$ ;
6     if (in transmit mode) then
7         | transmit a message containing  $\mathcal{A}(u)$  during each slot of the frame;
8     else
9         // in receive mode
10        foreach (clear message received during the frame) do
11            | let the received message be sent by node  $v$  containing set  $A$ ;
12            | add  $\langle v, A \cap \mathcal{A}(u) \rangle$  to the set of neighbors;
13        end foreach
14    end if
15 end while

```

Algorithm 5: Neighbor discovery algorithm for an asynchronous system with a bound on clock drift rate and knowledge of upper bound on maximum node degree.

range:

$$\frac{L}{1 + \delta} \leq \text{frame length in real-time} \leq \frac{L}{1 - \delta} \quad (9)$$

A node further divides each frame into three equal-sized slots—equal with respect to its local clock. Therefore the duration of a slot as measured by a node using its local clock is $\frac{L}{3}$. At the beginning of each frame, a node u randomly selects a channel from its available channel set, say c . It then transmits on c during each slot of the frame with probability $\min\left(\frac{1}{2}, \frac{|\mathcal{A}(u)|}{3\Delta_{est}}\right)$ and listens on c during the entire frame with the remaining probability. In the former case, u transmits the same message during each slot of the frame. In the latter case, u listens for any clear messages it may receive during *any part* of the frame. The partition of a frame into slots is not important for a receiving node; they are only used by a transmitting node. Also, a node may receive multiple clear messages while listening during a single frame. The execution of a node with respect to its local clock is shown in Fig. 1. The execution of the network with respect to common real time is shown in Fig. 2.

A formal description of the algorithm is shown in Algorithm 5. We now analyze the running time of the algorithm. In our analysis, unless otherwise stated, time refers to *real-time*. Of course, nodes do not have access to the real-time. We define two notions we use in our analysis as follows:

Definition 1 (aligned pair). We say that a pair of frames $\langle f, g \rangle$ is aligned if at least one slot of

f lies completely within g .

Definition 2 (overlapping frames). For a frame f and a node u , let $\text{overlap}(f, u)$ denote the subset of frames of u that overlap in real-time with f . Further, let $\text{overlap-all}(f)$ denote the subset of all frames that overlap in real-time with f .

For example, as per the execution shown in Fig. 2, pairs $\langle f_1, g_1 \rangle$ and $\langle f_2, h_1 \rangle$ are aligned whereas the pair $\langle f_1, h_1 \rangle$ is not. Also, $\text{overlap}(g_2, v) = \{f_1, f_2\}$ and $\text{overlap-all}(g_2) = \{f_1, f_2, g_2, h_1, h_2\}$.

One of the arguments we commonly use in our analysis is: “ x adjacent slots of a node cannot strictly contain y adjacent slots of another node” for certain specific values of x and y with $x < y$. To prove this, we argue that, for the statement to be false, the following inequality must hold:

$$\frac{xL}{3(1-\delta)} > \frac{yL}{3(1+\delta)}$$

and then show that it contradicts Assumption 1. The above inequality must hold for the statement to be false because the left hand side denotes the largest possible length of the time interval containing x adjacent slots, and the right hand side denotes the smallest possible length of the time interval containing y adjacent slots.

Lemma 5. A frame of a node overlaps with at most three frames of any other node. Formally,

$$\forall f, u :: |\text{overlap}(f, u)| \leq 3$$

Proof. Assume, in order to obtain a contradiction, that a frame of some node, say f , overlaps with at least four frames of another node. This implies that there are at least two consecutive frames that are strictly contained within f . From (9), a frame of a node can strictly contain two frames of another node only if the following condition holds:

$$\frac{L}{1-\delta} > \frac{2L}{1+\delta} \implies \delta > \frac{1}{3}$$

This contradicts Assumption 1. □

Consider the link from node v to node u on channel c . Also, consider frames f of v and g of u such that the pair $\langle f, g \rangle$ is aligned. We extend the notion of a link covered by a time-slot (defined

in Section 3.1.1) to a link covered by an aligned pair of frames. Specifically, the pair of aligned frames $\langle f, g \rangle$ covers the link (v, u) on channel c if: (i) v transmits on c during f , (ii) u listens on c during g , and (iii) no other neighbor of u , say w , transmits on c during any frame in $\text{overlap}(g, w)$. Also, $\langle f, g \rangle$ covers (v, u) if $\langle f, g \rangle$ covers (v, u) on some channel $c \in \text{span}(v, u)$. The three conditions collectively ensure that u receives a clear message from v (on c during g) provided $\langle f, g \rangle$ is aligned. We have:

Lemma 6. *If $\langle f, g \rangle$ is aligned, then $\langle f, g \rangle$ covers (v, u) with probability at least $\frac{\rho}{8 \max(2S, 3\Delta_{est})}$.*

Proof. Analogous to $A(\tau, c)$, $B(\tau, c)$ and $C(\tau, c)$ defined in Section 3.1.1, let $\hat{A}(f, g, c)$, $\hat{B}(f, g, c)$, $\hat{C}(f, g, c)$ denote the three events corresponding to the three conditions (i), (ii) and (iii), respectively, required for coverage. Note that the three events are *mutually independent*. As before, we compute the probability of occurrence of the three events separately.

Computing the probability of occurrence of $\hat{A}(f, g, c)$ We have:

$$\begin{aligned}
\Pr\{\hat{A}(f, g, c)\} &= (v \text{ selects } c \text{ at the beginning of } f) \wedge \\
&\quad (v \text{ chooses to transmit during } f) \\
&= \frac{1}{|\mathcal{A}(v)|} \times \min\left(\frac{1}{2}, \frac{|\mathcal{A}(v)|}{3\Delta_{est}}\right) \\
&= \min\left(\frac{1}{2|\mathcal{A}(v)|}, \frac{1}{3\Delta_{est}}\right) \\
&\geq \frac{1}{\max(2S, 3\Delta_{est})}
\end{aligned} \tag{10}$$

Computing the probability of occurrence of $\hat{B}(f, g, c)$ We have:

$$\begin{aligned}
\Pr\{\hat{B}(f, g, c)\} &= (u \text{ selects } c \text{ at the beginning of } g) \wedge \\
&\quad (u \text{ chooses to listen during } g) \\
&= \frac{1}{|\mathcal{A}(u)|} \times \left\{ 1 - \min\left(\frac{1}{2}, \frac{|\mathcal{A}(u)|}{3\Delta_{est}}\right) \right\} \\
&\quad \left\{ \min(x, y) \leq x \right\} \\
&\geq \frac{1}{|\mathcal{A}(u)|} \times \left(1 - \frac{1}{2}\right) = \frac{1}{2|\mathcal{A}(u)|}
\end{aligned} \tag{11}$$

Computing the probability of occurrence of $\hat{C}(f, g, c)$ Let $\mathcal{N}(u, c)$ denote the set of neighbors of u on c . Note that, if $\mathcal{N}(u, c)$ only contains v , then $\Pr\{\hat{C}(f, g, c)\} = 1$. Otherwise, we have:

$$\begin{aligned}
& \Pr\{\hat{C}(f, g, c)\} \\
&= \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \Pr \left(\begin{array}{c} w \text{ does not transmit on } c \text{ during any frame} \\ \text{in } \textit{overlap}(g, w) \end{array} \right) \\
&= \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \prod_{h \in \textit{overlap}(g, w)} \Pr \left(\begin{array}{c} w \text{ does not transmit on } c \\ \text{during } h \end{array} \right) \\
&= \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \prod_{h \in \textit{overlap}(g, w)} \{1 - \Pr(w \text{ transmits on } c \text{ during } h)\} \\
&\geq \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \prod_{h \in \textit{overlap}(g, w)} \left\{ 1 - \frac{1}{|\mathcal{A}(w)|} \times \min \left(\frac{1}{2}, \frac{|\mathcal{A}(w)|}{3\Delta_{est}} \right) \right\} \\
&= \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \prod_{h \in \textit{overlap}(g, w)} \left\{ 1 - \min \left(\frac{1}{2|\mathcal{A}(w)|}, \frac{1}{3\Delta_{est}} \right) \right\} \\
&\quad \left\{ \min(x, y) \leq y \right\} \\
&\geq \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \prod_{h \in \textit{overlap}(g, w)} \left(1 - \frac{1}{3\Delta_{est}} \right) \\
&\quad \left\{ \text{using Lemma 5, } |\textit{overlap}(g, w)| \leq 3 \right\} \\
&\geq \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \left(1 - \frac{1}{3\Delta_{est}} \right)^3 \\
&= \left(1 - \frac{1}{3\Delta_{est}} \right)^{3(|\mathcal{N}(u, c)| - 1)} \\
&\quad \left\{ |\mathcal{N}(u, c)| - 1 = \Delta(u, c) - 1 \leq \Delta_{est} \right\} \\
&\geq \left(1 - \frac{1}{3\Delta_{est}} \right)^{3\Delta_{est}} \\
&\quad \left\{ \forall x \geq 2, \left(1 - \frac{1}{x} \right)^x \text{ is a monotonically increasing} \right. \\
&\quad \left. \text{function of } x \text{ and thus } \geq \frac{1}{4} \right\}
\end{aligned}$$

$$\geq \frac{1}{4} \quad (12)$$

Finally, using a derivation similar to that for (6), it can be shown that $\langle f, g \rangle$ covers (v, u) with probability at least $\frac{\rho}{8 \max(2S, 3\Delta_{est})}$. \square

For a frame f , we use $node(f)$ to denote the node to which f belongs. For example, as per the execution show in Fig. 2, $node(f_1) = v$ and $node(h_1) = w$. We now define a precedence relation between pairs of frames referred to as *frame-pairs* for short.

Definition 3 (precedence relation). For frame-pairs $\langle f, g \rangle$ and $\langle p, q \rangle$, we say that $\langle f, g \rangle$ precedes $\langle p, q \rangle$, denoted $\langle f, g \rangle \sqsubset \langle p, q \rangle$, if (i) $node(f) = node(p)$, (ii) $node(g) = node(q)$, (iii) the start time of f is before that of p , and (iv) the start time of g is before that of q .

For example, as per the execution shown in Fig. 2, $\langle f_1, g_1 \rangle \sqsubset \langle f_2, g_2 \rangle$ and $\langle f_1, g_1 \rangle \sqsubset \langle f_3, g_2 \rangle$ but $\langle f_1, g_1 \rangle \not\sqsubset \langle f_1, g_2 \rangle$. We next define a notion on a sequence of frame-pairs that intuitively enables us to treat coverage provided by different frame-pairs as essentially independent events.

Definition 4 (admissible sequence). A sequence σ of M frame-pairs $\left\{ \langle f_{i_k}, g_{j_k} \rangle \right\}_{1 \leq k \leq M}$ is admissible with respect to the link (v, u) if it satisfies the following conditions:

1. $\forall i : 1 \leq k \leq M : node(f_{i_k}) = v$ and $node(g_{j_k}) = u$,
2. $\forall i : 1 \leq k < M : \langle f_{i_k}, g_{j_k} \rangle \sqsubset \langle f_{i_{k+1}}, g_{j_{k+1}} \rangle$,
3. $\forall i : 1 \leq k \leq M : \langle f_{i_k}, g_{j_k} \rangle$ is aligned, and
4. $\forall i : 1 \leq k < M : overlap-all(g_{i_k}) \cap overlap-all(g_{i_{k+1}}) = \emptyset$.

For a sequence σ of frame-pairs that is admissible with respect to (v, u) , we say that σ covers (v, u) if some frame-pair in the sequence covers (v, u) .

Lemma 7. Let σ be a sequence of $\frac{8 \max(2S, 3\Delta_{est})}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$ frame-pairs such that σ is admissible with respect to (v, u) . Then the probability that σ does not cover (v, u) is at most $\frac{\epsilon}{N^2}$.

Proof. For convenience, let $M = \frac{8 \max(2S, 3\Delta_{est})}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$. Let E_k with $1 \leq k \leq M$ denote the event that $\langle f_{i_k}, g_{j_k} \rangle$ covers (v, u) on some channel. Note that the occurrence of the event E_k

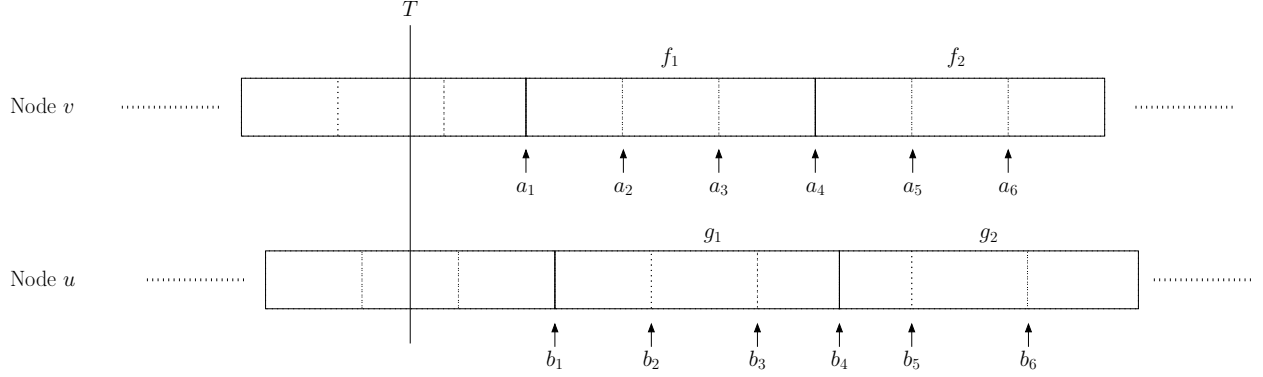


Figure 3: Slots used in the proof of Lemma 8.

depends *only* on frames in $overlap-all(g_{j_k})$. Since σ is an admissible sequence, for all x and y , with $1 \leq x < y \leq M$, $overlap-all(g_{j_x}) \cap overlap-all(g_{j_y}) = \emptyset$. In other words, the set of frames that overlap with the frame g_{j_x} are distinct from the set of frames that intersect the frame g_{j_y} . Since, for each frame, a node chooses its action randomly, events E_k s are *mutually independent* of each other. Using a derivation similar to that of (7), it can be shown that σ does not cover (v, u) with probability at most $\frac{\epsilon}{N^2}$. \square

We next show that any execution of the network must contain a “sufficiently long” sequence of admissible frame-pairs. In the remainder of this section, let T_s denote the time by which all nodes have initiated the neighbor discovery algorithm, Algorithm 5. To show the existence of a “sufficiently long” admissible sequence, we first show that, for any instant of time T after T_s and, for any pair of neighboring nodes, there is an aligned pair of frames after but “close” to T .

Lemma 8. *Consider a link from node v to node u and some instant of time T with $T \geq T_s$. Let f_i (respectively, g_i) with $i \geq 1$ denote the i^{th} full frame of node v (respectively, node u) after T . Then some frame in $\{f_1, f_2\}$ is aligned with some frame in $\{g_1, g_2\}$.*

Proof. Note that the frames f_1 and f_2 together contain six slots. Let a_i for $i = 1, 2, \dots, 6$ denote the start times of the six slots numbered in the increasing order of their start times (see Fig. 3). Likewise, let b_i for $i = 1, 2, \dots, 6$ denote the start times of the six slots of the frames g_1 and g_2 .

Claim 1: We first show that $b_1 \leq a_5$. Assume, by the way of contradiction, that $a_5 < b_1$. Note that, by definition of g_1 , there is only a partial frame of u between T and b_1 . This, in turn, implies

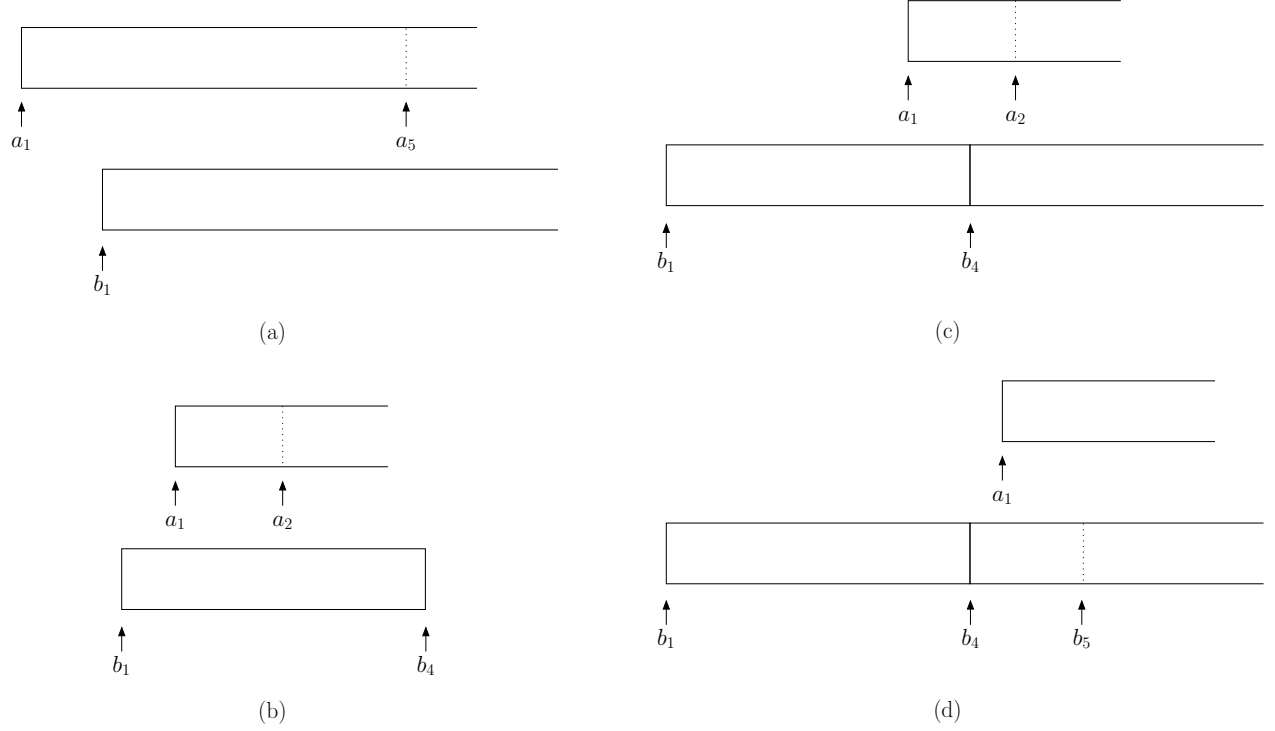


Figure 4: Various cases in the proof of Lemma 8.

that a partial frame of u contains at least four slots of v . This can happen only if the following condition holds:

$$\frac{L}{1-\delta} > \frac{4L}{3(1+\delta)} \implies \delta > \frac{1}{7}$$

This contradicts Assumption 1.

Claim 2: Likewise, we can show that $a_1 \leq b_5$.

Claim 3: If $a_{i-1} \leq b_1 \leq a_i$ for some i with $1 < i \leq 5$, then $b_1 \leq a_i < a_{i+1} \leq b_4$. It suffices to show that $a_{i+1} \leq b_4$. Assume, by the way of contradiction, that $b_4 < a_{i+1}$. This implies that $a_{i-1} \leq b_1 < b_4 < a_{i+1}$. In other words, two adjacent slots of v strictly contain an entire frame of u .

This can happen only if the following condition holds:

$$\frac{2L}{3(1-\delta)} > \frac{L}{1+\delta} \implies \delta > \frac{1}{5}$$

This contradicts Assumption 1.

We now prove the lemma statement. We consider two cases depending on whether f_1 or g_1 has earlier start time.

- *Case 1* ($a_1 < b_1$): In this case, we have $a_1 < b_1 \leq a_5$ (see Fig. 4(a)). The rightmost inequality follows from Claim 1. The time interval from a_1 to a_5 consists of four contiguous slots of v : $[a_1, a_2]$, $[a_2, a_3]$, $[a_3, a_4]$ and $[a_4, a_5]$. Clearly, b_1 lies in at least one of them. (If b_1 lies on the boundary of two slots, we select the earlier one.) In any case, using Claim 3, we can show that at least one of the slots $[a_2, a_3]$, $[a_3, a_4]$, $[a_4, a_5]$ or $[a_5, a_6]$ is contained in the frame g_2 . This implies that either $\langle f_1, g_1 \rangle$ or $\langle f_2, g_1 \rangle$ is aligned.
- *Case 2* ($b_1 \leq a_1$): If $a_2 \leq b_4$, then $b_1 \leq a_1 < a_2 \leq b_4$ (see Fig. 4(b)). In other words, the first slot of the frame f_1 is contained within the frame g_1 , which implies that $\langle f_1, g_1 \rangle$ is aligned. Therefore assume that $b_4 < a_2$. We have two subcases depending on where a_1 lies relative to b_4 . Let b_7 denote the end time of the frame g_2 .
 - *Case 2.1* ($a_1 \leq b_4$): In this case, we show that the slot $[a_2, a_3]$ is contained within the frame $[b_4, b_7]$ (see Fig. 4(c)). To that end, we first prove that $a_3 < b_7$. If not, then $a_1 \leq b_4 < b_7 \leq a_3$. In other words, two adjacent slots of v strictly contain an entire frame of u . This can happen only if the following condition holds:

$$\frac{2L}{3(1-\delta)} > \frac{L}{1+\delta} \implies \delta > \frac{1}{5}$$

This contradicts Assumption 1. Therefore, we have $b_4 < a_2 < a_3 < b_7$, which implies that the pair $\langle f_1, g_2 \rangle$ is aligned.

- *Case 2.2* ($b_4 < a_1$): In this case, we show that the slot $[a_1, a_2]$ is contained within the frame $[b_4, b_7]$ (see Fig. 4(d)). To that end, we show that $a_2 \leq b_7$. If not, then $a_1 \leq b_5 < b_7 < a_2$. (The inequality $a_1 \leq b_5$ follows from Claim 2.) In other words, one slot of v strictly contains at least two adjacent slots of u . This can happen only if the following condition holds:

$$\frac{L}{3(1-\delta)} > \frac{2L}{3(1+\delta)} \implies \delta > \frac{1}{3}$$

This contradicts Assumption 1. Therefore, we have $b_4 < a_1 < a_2 \leq b_7$, which implies that the pair $\langle f_1, g_2 \rangle$ is aligned.

In all cases, we show that one of the four pairs $\langle f_1, g_1 \rangle$, $\langle f_1, g_2 \rangle$, $\langle f_2, g_1 \rangle$ or $\langle f_2, g_2 \rangle$ is aligned. \square

We now show that existence of a “sufficiently long” admissible sequence in any execution of Algorithm 5.

Lemma 9. *Consider a link from node v to node u . Further, consider an execution of the network after T_s that contains at least M full frames of u as well as v . Then the execution contains a sequence of at least $\frac{M}{6}$ frame-pairs such that the sequence is admissible with respect to (v, u) .*

Proof. The proof is by construction. The construction is in two steps. In the first step, we construct a sequence of frame-pairs γ that is “almost” admissible in the sense that it satisfies the first three properties of an admissible sequence but may not satisfy the fourth property. We show that γ contains at least $\frac{M}{2}$ frame-pairs. In the second step, using γ , we construct a sequence of frame-pairs σ that satisfies all four properties of an admissible sequence. We also show that σ contains at least $\frac{M}{6}$ frame-pairs.

Constructing γ To obtain the first frame-pair $\langle f_{i_1}, g_{i_1} \rangle$ that is aligned, we apply Lemma 8 to T_s . Now, assume that we have already selected k frame-pairs satisfying the first three properties of an admissible sequence. Let the k^{th} frame-pair be denoted by $\langle f_{i_k}, g_{i_k} \rangle$. To select the next frame-pair, let T_k be defined as the *earlier* of the end times of frames f_{i_k} and g_{i_k} . To obtain the next frame-pair that is aligned, we apply Lemma 8 to T_k . Let the frame-pair be denoted by $\langle f_{i_{k+1}}, g_{i_{k+1}} \rangle$. Clearly, the extended sequence (consisting of $k+1$ aligned frame-pairs) satisfies the first and third properties of an admissible sequence. The second property holds because the start times of both f_{i_k} and g_{i_k} are *before* T_k , whereas the start times of both $f_{i_{k+1}}$ and $g_{i_{k+1}}$ are *after* T_k . We repeatedly select aligned frame-pairs using Lemma 8 until we run out of frames of either u or v . We now establish a lower bound on the length of γ . Note that, when selecting the $(k+1)^{st}$ pair, the first full frame of node v after T_k (namely frame f_1 in the statement of Lemma 8) is *adjacent* to the frame f_{i_k} . This follows from the definition of T_k . This, in turn, implies that the frame $f_{i_{k+1}}$ obtained using Lemma 8 is within a distance of two of the frame f_{i_k} . Likewise, the frame $g_{i_{k+1}}$ is within a distance of two of the frame g_{i_k} . As a result, γ contains at least $\frac{M}{2}$ frame-pairs.

Constructing σ To construct a sequence σ that also satisfies the fourth property of an admissible sequence, we choose every *third* frame-pair of γ starting with the first frame-pair $\langle f_{i_1}, g_{i_1} \rangle$. Clearly σ also satisfies the first three properties of an admissible sequence (since it is a subsequence of γ). Let the k^{th} frame-pair of σ be denoted by $\langle f_{j_k}, g_{j_k} \rangle$. To prove that σ satisfies the fourth property as well, assume, by the way of contradiction, that some frame, say h , overlaps with two consecutive frames g_{j_k} and $g_{j_{k+1}}$ for some k . Note that, since we selected only every third frame-pair of γ to construct σ , there are at least two other frames of u between g_{j_k} and $g_{j_{k+1}}$. This implies that h overlaps with at least four frames of u , which contradicts Lemma 5. This establishes that the length of σ is at least $\frac{M}{6}$. \square

Finally, we have the main result.

Theorem 10. *Let T_s be the time by which all nodes have initiated neighbor discovery. Also, let T_f be the earliest time by which each node has executed at least $\frac{48 \max(2S, 3\Delta_{est})}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$ full frames since T_s . Then Algorithm 5 ensures that each node discovers all its neighbors on all channels with probability at least $1 - \epsilon$ by time T_f .*

Proof. Consider a link (v, u) . From Lemma 9, the execution from T_s to T_f contains a sequence of at least $\frac{8 \max(2S, 3\Delta_{est})}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$ frame-pairs such that the sequence is admissible with respect to (v, u) . From Lemma 7, the probability that (v, u) is not covered by T_f is at most $\frac{\epsilon}{N^2}$. This implies that the probability that some link in the network is not covered by T_f is at most ϵ . \square

We now bound the length of the interval $T_f - T_s$.

Theorem 11. *Let T_s and T_f be as defined in Theorem 10. Then the length of the interval $T_f - T_s$ is upper bounded by $\left\{ \frac{48 \max(2S, 3\Delta_{est})}{\rho} \ln \left(\frac{N^2}{\epsilon} \right) + 1 \right\} \left(\frac{L}{1-\delta} \right)$.*

Proof. By our choice of T_f , there exists some node such that $T_f - T_s$ contains exactly $\frac{48 \max(2S, 3\Delta_{est})}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$ full frames of that node. The first frame of that node in the execution may be partial frame. From (9), the length of each frame is upper bounded by $\frac{L}{1-\delta}$. Combining the three facts, we obtain the result. \square

In our algorithm, we divided our frame into three slots. It is possible to divide a frame into only two slots and a result similar to Theorem 11 can be derived. However, the proof is more involved.

```

// Algorithm for node  $u$ 
Output: The set of neighbors along with the subset of channels that are common with the neighbor.

1  $i \leftarrow 1$ ;
2 while true do
    // epoch  $i$ 
3   for  $j \leftarrow 1$  to  $i + 1$  do
        // compute the transmission probability
4        $p \leftarrow \min\left(\frac{1}{2}, \frac{|\mathcal{A}(u)|}{3 \cdot 2^j}\right)$ ;
5       for  $k \leftarrow 1$  to  $2^i$  do
            // phase  $j$  in epoch  $i$ 
            // select a channel
6            $c \leftarrow$  channel selected uniformly at random from  $\mathcal{A}(u)$ ;
7           tune the transceiver to  $c$ ;
8           switch to transmit mode with probability  $p$  and receive mode with probability  $1 - p$ ;
9           if (in transmit mode) then
10              transmit a message containing  $\mathcal{A}(u)$  during each slot of the frame;
11          else
                // in receive mode
12              foreach (clear message received during the frame) do
13                  let the received message be sent by node  $v$  containing set  $A$ ;
14                  add  $\langle v, A \cap \mathcal{A}(u) \rangle$  to the set of neighbors;
15              end foreach
16          end if
17      end for
18  end for
19   $i \leftarrow i + 1$ ;
20 end while

```

Algorithm 6: Neighbor discovery algorithm for an asynchronous system with a bound on clock drift rate.

Likewise, a frame can be divided into four or more slots albeit at the expense of increasing the running time of the algorithm.

4.2 No Knowledge of Maximum Node Degree

The algorithm for this case is similar to the one for a synchronous system, namely Algorithm 4, but with time-slots replaced with frames and transmission probability calculated as in Algorithm 5. A formal description of the algorithm is given in Algorithm 6.

The time-complexity analysis of Algorithm 6 is more involved because of asynchrony. Unlike in a synchronous system, the time difference between when two nodes start the same phase may *grow* with time due to clock drift. We need some additional notation for the analysis.

Let $\Lambda(i, j)$ denote the phase j of the epoch i . Let $F_{i,j}$ denote the number of frames that a node

in the execution until (but not including) $\Lambda(i, j)$. We have:

$$\begin{aligned}
F_{i,j} &= (\text{number of frames contained in the first } i-1 \text{ epochs}) + \\
&\quad (\text{number of frames contained in the } i^{\text{th}} \text{ epoch until } \Lambda(i, j)) \\
&\quad \{ \text{using (A.6) in the appendix and the fact that, for each } j, |\Lambda(i, j)| = 2^i \} \\
&= (i-1) \cdot 2^i + (j-1) \cdot 2^i \\
&= (i+j-2) \cdot 2^i
\end{aligned} \tag{13}$$

Let $T_b^u(i, j)$ denote the time at which node u begins executing $\Lambda(i, j)$. Likewise, let $T_e^u(i, j)$ denote the time at which node u ends executing $\Lambda(i, j)$. Let $T_b^{\max}(i, j)$ denote the *latest* time at which some node starts executing $\Lambda(i, j)$. Likewise, let $T_e^{\min}(i, j)$ denote the *earliest* time at which some node starts executing $\Lambda(i, j)$. Symbols T_s , Θ and L have the same meaning as before. Note that, in real-time, the minimum and maximum lengths of a frame are given by $\frac{L}{1+\delta}$ and $\frac{L}{1-\delta}$, respectively. Thus, we have:

$$T_b^{\max}(i, j) = \max\{T_b^u(i, j)\} \leq T_s + F_{i,j} \left(\frac{L}{1-\delta} \right) \tag{14}$$

$$T_e^{\min}(i, j) = \min\{T_e^u(i, j)\} \geq (T_s - \Theta) + (F_{i,j} + 2^i) \left(\frac{L}{1+\delta} \right) \tag{15}$$

Suppose i and j are such that $T_e^{\min}(i, j) < T_b^{\max}(i, j)$. Note that, during the time-period from $T_b^{\max}(i, j)$ to $T_e^{\min}(i, j)$, all nodes will use the *same* estimate for maximum node degree. Recall that Δ_0 denotes the smallest power of two greater than or equal to δ . Let $M = \frac{48 \max(2S, 3\Delta_0)}{\rho} \ln \left(\frac{N^2}{\epsilon} \right) + 1$. From Theorem 10, if the time period from $T_b^{\max}(i, j)$ to $T_e^{\min}(i, j)$ is such that: (a) the time period contains at least M full frames and (b) each node uses an estimate value of Δ_0 during the time period, then the neighbor discovery is guaranteed to complete with probability at least $1 - \epsilon$. Let $i_0 = \lceil \log(6M + \frac{3\Theta}{L}) \rceil$ and $j_0 = \lceil \log \Delta \rceil (= \log \Delta_0)$. Note that, by definition, during $\Lambda(i_0, j_0)$, each node uses Δ_0 as the estimate for maximum node degree. It suffices to prove that the time period from $T_s i_0 j_0$ to $T_f i_0 j_0$ contains at least M frames. To that end, we make the following assumption about the clock drift rate:

Assumption 2. Let $D = 48 \left\{ \log \left(\frac{NS}{\epsilon} \right) + \log \left(1 + \frac{\Theta}{L} \right) + 5 \right\}$. The maximum drift rate of the clock

of any node is bounded by $\frac{1}{D}$ seconds/second.

We first prove the following lemma:

Lemma 12. *If Assumption 2 holds, then $i_0 \cdot \delta \leq \frac{1}{16}$.*

Proof. We have:

$$\begin{aligned} M &= \frac{48 \max(2S, 3\Delta_0)}{\rho} \ln \left(\frac{N^2}{\epsilon} \right) + 1 \\ &\leq \left(\frac{48 \cdot (2S) \cdot (6N)}{\frac{1}{S}} \right) \cdot \left(\frac{N^2}{\epsilon} \right) + 1 \\ &= \frac{576 \cdot S^2 \cdot N^3}{\epsilon} + 1 \end{aligned}$$

This implies that:

$$\begin{aligned} \log M &\leq \log 576 + 3 \log \left(\frac{NS}{\epsilon} \right) + 1 \\ &\leq 3 \log \left(\frac{NS}{\epsilon} \right) + 11 \end{aligned} \tag{16}$$

Now, we have:

$$\begin{aligned} i_0 \cdot \delta &\leq \left\{ \log \left(6M + \frac{3\Theta}{L} \right) + 1 \right\} \cdot \delta \\ &\leq \left\{ \log 6 + \log \left(M + \frac{\Theta}{L} + 1 \right) + 1 \right\} \cdot \delta \\ &\quad \{\text{for all } x, y \text{ with } x \geq 1 \text{ and } y \geq 1, \log(x+y) \leq \log(xy) \leq \log x + \log y\} \\ &\leq \left\{ \log(M) + \log \left(1 + \frac{\Theta}{L} \right) + 4 \right\} \cdot \delta \\ &\quad \{\text{using (16)}\} \\ &\leq \left\{ 3 \log \left(\frac{NS}{\epsilon} \right) + 11 + \log \left(1 + \frac{\Theta}{L} \right) + 4 \right\} \cdot \delta \\ &\leq 3 \cdot \left\{ \log \left(\frac{NS}{\epsilon} \right) + \log \left(1 + \frac{\Theta}{L} \right) + 5 \right\} \cdot \delta \\ &\quad \{\text{using definition of } D\} \\ &\leq 3 \cdot \frac{D}{48} \cdot \delta \\ &\quad \{\text{using Assumption 2}\} \end{aligned}$$

$$\leq 3 \cdot \frac{D}{48} \cdot \frac{1}{D} = \frac{1}{16}$$

This establishes the lemma. \square

Now, substituting $i = i_0$ and $j = j_0$ in (14) and (15) and subtracting (14) from (15), we obtain:

$$\begin{aligned}
& T_e^{\min}(i_0, j_0) - T_b^{\max}(i_0, j_0) \\
& \geq (F_{i_0, j_0} + 2^{i_0}) \left(\frac{L}{1 + \delta} \right) - F_{i_0, j_0} \left(\frac{L}{1 - \delta} \right) - \Theta \\
& \quad \{\text{rearranging terms}\} \\
& = 2^{i_0} \left(\frac{L}{1 + \delta} \right) - F_{i_0, j_0} \left(\frac{2\delta L}{1 - \delta^2} \right) - \Theta \\
& \quad \{\text{using (13)}\} \\
& = 2^{i_0} \left(\frac{L}{1 + \delta} \right) - (i_0 + j_0 - 2) \cdot 2^{i_0} \left(\frac{2\delta L}{1 - \delta^2} \right) - \Theta \\
& \geq 2^{i_0} \left(\frac{L}{1 + \delta} \right) \left\{ 1 - (i_0 + j_0 - 2) \left(\frac{2\delta}{1 - \delta} \right) \right\} - \Theta \\
& \quad \{i_0 + j_0 - 2 \leq 2i_0\} \\
& \geq 2^{i_0} \left(\frac{L}{1 + \delta} \right) \left(1 - \frac{(2i_0)(2\delta)}{1 - \delta} \right) - \Theta \\
& \quad \left\{ \delta \leq \frac{1}{2} \right\} \\
& \geq 2^{i_0} \left(\frac{2L}{3} \right) (1 - 8 \cdot i_0 \cdot \delta) - \Theta \\
& \quad \{\text{using Lemma 12}\} \\
& \geq 2^{i_0} \left(\frac{L}{3} \right) - \Theta \\
& \geq \left(6M + \frac{3\Theta}{L} \right) \left(\frac{L}{3} \right) - \Theta \\
& = 2ML + \Theta - \Theta = 2ML
\end{aligned}$$

We now prove the main result. Since the maximum length of a frame in real-time is $\frac{L}{1-\delta}$ and $\delta \leq \frac{1}{2}$, it follows that the time-period from $T_b^{\max}(i_0, j_0)$ to $T_b^{\max}(i_0, j_0)$ contains at least $\frac{2ML}{L/1-\delta} = 2M(1-\delta) \geq M$ full frames.

The above analysis implies that, once all nodes have completed their epoch i_0 , the neighbor

Table 1: A summary of all randomized neighbor discovery algorithms in this work.

Assumptions			Time Complexity*	Section
Synchrony	Start Times	System Parameters		
Synchronous	Same	know an estimate for an upper bound on Δ , given by Δ_{est}	$O\left(\frac{\max(S, \Delta)}{\rho} \log(\Delta_{est}) \log\left(\frac{N}{\epsilon}\right)\right)$	3.1.1
Synchronous	Same	-	$O(M \log M)$ where $M = \frac{16 \max(S, \Delta)}{\rho} \ln\left(\frac{N^2}{\epsilon}\right)$	3.1.2
Synchronous	Different	know an estimate for an upper bound on Δ , given by Δ_{est}	$O\left(\frac{\max(2S, \Delta_{est})}{\rho} \log\left(\frac{N}{\epsilon}\right)\right)$	3.2.1
Synchronous	Different	-	$O((M + \Theta) \log(M + \Theta))$ where $M = \frac{16 \max(S, \Delta_0)}{\rho} \ln\left(\frac{N^2}{\epsilon}\right)$	3.2.2
Asynchronous	Different	know an estimate for an upper bound on Δ , given by Δ_{est} , and $\delta \leq \frac{1}{7}$	$O\left(\frac{L \max(2S, 3\Delta_{est})}{\rho} \log\left(\frac{N}{\epsilon}\right)\right)$	4.1
Asynchronous	Different	$\delta \leq \frac{1}{48\{\log(\frac{NS}{\epsilon}) + \log(1 + \frac{\Theta}{L}) + 5\}}$	$O\left((ML + \Theta) \log\left(M + \frac{\Theta}{L}\right)\right)$ where $M = \frac{48 \max(2S, 3\Delta_0)}{\rho} \ln\left(\frac{N^2}{\epsilon}\right) + 1$	4.2

*: In the case of variable start times, the time is measured after the last node has started neighbor discovery

N : total number of CR nodes in the network

Δ : maximum degree of a node on a channel

S : size of the largest available channel set

ϵ : maximum error probability (user-specified)

Θ : maximum difference in the start times of two nodes

L : length of a frame

Δ_0 : smallest power of two greater than or equal to Δ
 ρ : minimum span ratio of a link

δ : maximum drift rate of a local clock

discovery completes with probability at least $1 - \epsilon$. Thus we have:

Theorem 13. *Let $M = \frac{48 \max(2S, 3\Delta_0)}{\rho} \ln\left(\frac{N^2}{\epsilon}\right) + 1$. Then Algorithm 6 ensures that each node discovers all its neighbors on all channels within $O\left((M + \frac{\Theta}{L}) \log\left(M + \frac{\Theta}{L}\right) \left(\frac{L}{1-\delta}\right)\right)$ time-units of T_s with probability at least $1 - \epsilon$.*

Remark (Impact of Assumption 2). *Even for very large values of system parameters, namely $N = 10^6$, $S = 10^6$, $\epsilon = 10^{-9}$ and $\frac{\Theta}{L} = 10^6$, Assumption 2 only requires $\delta \leq 2.2 \times 10^{-3}$ seconds/seconds. This is not very restrictive because, as we mentioned earlier, an ordinary quartz clock has a maximum drift rate of 10^{-6} seconds/second.*

5 Summary

In this paper, we have presented several randomized algorithms for neighbor discovery in an M²HeW network both for a synchronous and an asynchronous system, and also analyzed their running time. Table 1 summarizes the results of the previous two sections. We now discuss two extensions to our algorithms to enhance their applicability.

6 Improving Robustness

6.1 Handling Diverse Propagation Characteristics

All algorithms described so far have assumed “one-for-all” property that if a link from node v to node u exists on some channel then it also exists on all channels that are present in available channel sets of both u and v . That may not hold if different channels have very different propagation characteristics. In case the “one-for-all” property does not hold, we can partition the channels into *bands* such that the subset of channels within the same band have very similar propagation characteristics. As a result, “one-for-all” holds for each band individually. We say that a link can operate in a band if it can operate on some channel in the band. Each link can then be replaced with multiple links, one for each band it can operate in. Link for each band then has to be discovered separately. This modification increases the running time of our algorithms since more links have to be discovered now. The minimum span-ratio in the network now ranges from $\frac{1}{S}$ to $\frac{W}{S}$ instead of $\frac{1}{S}$ to 1, where W is the size of the smallest band (in terms of number of channels). Further the term “ $\ln\left(\frac{N^2}{\epsilon}\right)$ ” in time complexity expressions is replaced with the term “ $\ln\left(\frac{N^2 B}{\epsilon}\right)$ ”, where B denote the maximum number of bands in which any link can operate.

6.2 Handling Lossy Channels

All algorithms described so far have assumed that channels are reliable in the sense that a node fails to receive a message from its neighbor only due to collision. However, in the real world, a node may fail to receive a message due to other reasons as well such as Gaussian noise or transient interference. Let ϕ with $0 \leq \phi < 1$ denote the probability that a node fails to receive a message from its neighbor even in the absence of collision. Nodes may not know ϕ . With lossy channels, the probability that a link is covered by a time-slot (in Algorithm 1, Algorithm 2, Algorithm 3 and Algorithm 4) or an aligned frame-pair (in Algorithm 5 and Algorithm 6) is now multiplied by a factor of $1 - \phi$. For example, the probability expression in Lemma 6 now becomes $\frac{\rho(1-\phi)}{8 \max(2S, 3\Delta_{est})}$. It can be verified that this has the effect of increasing the running time of an algorithm by a factor of $\frac{1}{1-\phi}$.

6.3 Handling Jamming Attacks

Most of the work on jamming attack [18, 23, 38, 59] is concerned with analyzing or mitigating the impact of jamming attack on communication throughput in the system, and, as such, they assume that neighbor discovery has already been done. Work in [6] presents a suite of neighbor discovery algorithms in the presence of jamming attack. However, unlike our work, the work in [6] assumes a single-hop network and does not provide any theoretical guarantees (algorithms are only evaluated using simulation experiments).

We show that, under certain conditions, our neighbor discovery algorithms, with minor modifications, are tolerant to jamming attacks in the sense that a malicious jammer who seeks to prevent neighbor discovery can only delay time to the completion of neighbor discovery by at most a constant factor.

In formulating the jamming model, we seek a balance between a jammer who is too powerful (making the neighbor discovery problem impossible to solve) and one who is too weak (making the attack easy to foil). For instance, if the jammer can simultaneously jam all channels at all times, then neighbor discovery would be unable to complete. Likewise, if for some pair of neighboring nodes, $|span(u, v)| = 1$, then the jammer can ensure that discovery does not complete by simply jamming this channel.

We assume an “intelligent” jammer who can *scan* the spectrum and learn which channels are currently under use by transmitting nodes running the neighbor discovery algorithm. The jammer can then decide to jam any single channel, thereby disrupting possibly all transmissions on that channel. When scanning, the jammer only learns which channels are being used; we assume that it cannot determine which channels are used by a majority of nodes by analyzing the relative strength of the transmissions.

In a heterogeneous system, it may be possible for a jammer to increase its effectiveness by keeping statistics on which channels are seen in scans more often. In some cases, this would correspond to channels used by more nodes. It is not clear how best to analyze this case, so we explicitly avoid it by directing our analysis towards homogeneous systems in which $|span(u, v)| = S$ for all neighboring nodes u and v .

The jammer can change the channel which it jams during the course of the algorithm. Clearly,

if the jammer is able to switch channels instantaneously, it would be possible for the jammer to block every channel in every slot making it as powerful as the one who can simultaneously jam all channels at all times. To avoid this, we assume a lower bound on the time between switching channels for jamming. In the asynchronous case, this lower bound also allows to assume that the outcomes of two consecutive scans are independent of each other.

Note that the jammer need not respect slot boundaries even when the system is synchronous. Therefore the jammer can effectively jam *two channels* in a given slot by selecting a new channel to jam near the middle of a slot.

We also assume that the jammer has complete knowledge of our algorithm, but is ignorant of the exact random choices which are made. Clearly, the jammer's best option is to make a random selection of channels to jam. However, since the system is homogeneous, the best that the jammer can do is make a selection at random, in an attempt to foil the random choices made by our algorithm.

Based on the discussion above, we now formulate our jamming model. The jammer operates in rounds. At the beginning of each round, the jammer scans the spectrum and selects a channel to jam during that round from among the channels it observed to be in use in the scan. Specifically, let j denote the channel jammed by the jammer in the previous round. Also, let C denote the set of channels the jammer found to be in use in the scan it performed at the beginning of the current round. The jammer then makes a uniform random selection of a channel from the set $C \setminus j$, if non-empty, and jams this channel during the current round. We assume that round length is at least as large as slot length, but slot boundaries and round boundaries need not be aligned even if the system is synchronous. (A jamming round can start in the middle of a slot.) Note that it is not necessary for the jammer to jam j again since it would have already disrupted all transmissions on channel j in the current slot.

Note that, in the above-described jamming model, the neighbor discovery problem can only be solved if there are at least three nodes and at least three channels. We now consider synchronous and asynchronous cases separately.

6.3.1 Synchronous Algorithms

Recall the events $A(\tau, c)$, $B(\tau, c)$ and $C(\tau, c)$ defined earlier when analyzing the probability that node u discovers node v on channel c during time-slot τ . We define a new event $D(\tau, c)$ to represent the condition that the transmission from node v to node u on channel c is not jammed during the time-slot τ . For convenience, let $ABC(\tau, c)$ denote the conjunction of the three events $A(\tau, c)$, $B(\tau, c)$ and $C(\tau, c)$. We derive a lower bound on the conditional probability that event $D(\tau, c)$ occurs given that event $ABC(\tau, c)$ occurs. Note that the *slow-down factor* in the running time of a neighbor discovery algorithm that is undergoing a jamming attack is given by the *inverse of this conditional probability*.

Assume that, at the beginning of the time-slot τ , the jammer was jamming channel j . Sometime during τ , the jammer switches to a new channel to jam. We have:

$$\Pr\{D(\tau, c) \mid ABC(\tau, c)\} = \left(\frac{S-1}{S}\right) \cdot \Pr\{E(\tau, c) \mid ABC(\tau, c)\} \cdot \left(\frac{1}{2}\right) \quad (17)$$

where $E(\tau, c)$ denotes the event that some node in the system transmits on a channel other than c and j during the time-slot τ with $c \neq j$. In the above expression, the first term (given by $\frac{S-1}{S}$) captures the probability that $c \neq j$ and the third term (given by $\frac{1}{2}$) the probability that the jammer chooses a channel other than c to jam.

We next compute the conditional probability that the event $E(\tau, c)$ occurs given that the event $ABC(\tau, c)$ occurs. We partition the set of nodes in the system into three mutually-exclusive groups: G_1 , G_2 and G_3 . Group G_1 contains nodes u and v , group G_2 contains nodes that are neighbors of node u except for node v , and group G_3 contains the remaining nodes.

Let \mathcal{A} denote the set of all available channels. (Note that, by our assumption, all nodes have identical available channel sets.) Also, let $F(\tau, c, w)$ denote the event that node w *does not* transmit on a channel in $\mathcal{A} \setminus \{c, j\}$ during time-slot τ . Since nodes act independently of each other, we have:

$$\Pr\{E(\tau, c) \mid ABC(\tau, c)\} = 1 - \prod_{i \in \{1,2,3\}} \prod_{w \in G_i} \Pr\{F(\tau, c, w) \mid ABC(\tau, c)\} \\ \left\{ \forall w : w \in G_1 : \Pr\{F(\tau, c, w) \mid ABC(\tau, c)\} = 1 \right\}$$

$$\begin{aligned}
&= 1 - \prod_{i \in \{2,3\}} \prod_{w \in G_i} \Pr\{F(\tau, c, w) \mid ABC(\tau, c)\} \\
&\quad \left\{ \begin{array}{l} \forall w : w \in G_3 : \Pr\{F(\tau, c, w) \mid ABC(\tau, c)\} = \Pr\{F(\tau, c, w)\} \\ \because \text{ nodes act independently} \end{array} \right\} \\
&= 1 - \left(\prod_{w \in G_2} \Pr\{F(\tau, c, w) \mid ABC(\tau, c)\} \right) \cdot \left(\prod_{w \in G_3} \Pr\{F(\tau, c, w)\} \right) \quad (18)
\end{aligned}$$

Let p denote the placeholder for transmission probability. Note that $p = \min\left(\frac{1}{2}, \frac{S}{\Delta_{est}}\right)$. For our analysis, we focus on time-slots for which Δ_{est} satisfies the following condition:

$$\Delta_{est} \leq 2(N - 2) \quad (19)$$

We show that the above-condition trivially holds for three out of four synchronous algorithms, namely Algorithm 1, Algorithm 2 and Algorithm 4. For Algorithm 3, in which Δ_{est} is an “external” parameter, we state it as an explicit assumption. In the three synchronous algorithms (in which nodes estimate degree), for the time-slots used to show that neighbor discovery completes with arbitrarily high probability, note that Δ_{est} is the smallest power of two that is greater than or equal to either $\Delta(u, c)$ or Δ . In either case, we have:

$$\begin{aligned}
&\frac{\Delta_{est}}{2} < \Delta \\
\implies \frac{\Delta_{est}}{2} &\leq \Delta - 1 \\
&\left\{ \Delta \leq N - 1 \right\} \\
\implies \frac{\Delta_{est}}{2} &\leq (N - 1) - 1 = N - 2
\end{aligned}$$

We now derive the conditional probability $\Pr\{F(\tau, c, w) \mid ABC(\tau, c)\}$. There are two cases depending on whether $w \in G_2$ or $w \in G_3$. First, assume that $w \in G_2$. We have:

$$\begin{aligned}
\Pr\{F(\tau, c, w) \mid ABC(\tau, c)\} &= \Pr\{w \text{ does not transmit on } \mathcal{A} \setminus \{c, j\} \mid w \text{ does not transmit on } c\} \\
&= \frac{\Pr\{X \mid Y\} = \frac{\Pr\{X \wedge Y\}}{\Pr\{Y\}}}{\Pr\{(w \text{ does not transmit on } \mathcal{A} \setminus \{c, j\}) \wedge (w \text{ does not transmit on } c)\}} \\
&= \frac{\Pr\{X \wedge Y\}}{\Pr\{Y\}}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\Pr\{(w \text{ listens}) \text{ or } (w \text{ transmits on channel } j)\}}{\Pr\{(w \text{ listens}) \text{ or } (w \text{ transmits on a channel other than } c)\}} \\
&= \frac{1 - p + p \cdot \frac{1}{S}}{1 - p + p \cdot \frac{S-1}{S}} \\
&= \frac{1 - p \frac{S-1}{S}}{1 - \frac{p}{S}} \\
&\quad \left\{ \text{multiplying numerator and denominator by } \frac{S}{p} \right\} \\
&= \frac{\frac{S}{p} - S + 1}{\frac{S}{p} - 1} \\
&= 1 - \frac{S-2}{\frac{S}{p} - 1}
\end{aligned} \tag{20}$$

Next, assume that $w \in G_3$. We have:

$$\begin{aligned}
\Pr\{F(\tau, c, w) \mid ABC(\tau, c)\} &= \Pr\{w \text{ does not transmit on } \mathcal{A} \setminus \{c, j\}\} \\
&= \Pr\{(w \text{ listens}) \text{ or } (w \text{ transmits on channel } c \text{ or } j)\} \\
&= 1 - p + p \cdot \frac{2}{S} \\
&= 1 - p \cdot \frac{S-2}{S} \\
&= 1 - \frac{S-2}{\frac{S}{p}}
\end{aligned} \tag{21}$$

Using (18), (20) and (21), we have:

$$\begin{aligned}
\Pr\{E(\tau, c) \mid ABC(\tau, c)\} &= 1 - \prod_{w \in G_2} \left(1 - \frac{S-2}{\frac{S}{p} - 1}\right) \cdot \prod_{w \in G_3} \left(1 - \frac{S-2}{\frac{S}{p}}\right) \\
&\geq 1 - \prod_{w \in G_2} \left(1 - \frac{S-2}{\frac{S}{p}}\right) \cdot \prod_{w \in G_3} \left(1 - \frac{S-2}{\frac{S}{p}}\right) \\
&= 1 - \prod_{w \in G_2 \cup G_3} \left(1 - \frac{S-2}{\frac{S}{p}}\right) \\
&= 1 - \left(1 - p \cdot \frac{S-2}{S}\right)^{N-2}
\end{aligned} \tag{22}$$

We analyze the cases corresponding to when $p = \frac{1}{2}$ and when $p = \frac{S}{\Delta_{est}}$ separately.

- **Case 1** ($p = \frac{1}{2}$): Using (17) and (22), we have:

$$\begin{aligned}
\Pr\{D(\tau, c) \mid ABC(\tau, c)\} &\geq \frac{1}{2} \cdot \left(\frac{S-1}{S}\right) \cdot \left\{1 - \left(1 - \frac{S-2}{2S}\right)^{N-2}\right\} \\
&\quad \left\{N \geq 3\right\} \\
&\geq \frac{1}{2} \cdot \left(\frac{S-1}{S}\right) \cdot \left(\frac{S-2}{2S}\right) \\
&= \frac{1}{4} \cdot \left(1 - \frac{1}{S}\right) \cdot \left(1 - \frac{2}{S}\right) \\
&\quad \left\{S \geq 3\right\} \\
&\geq \frac{1}{4} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{1}{18}
\end{aligned}$$

- **Case 2** ($p = \frac{S}{\Delta_{est}}$): Using (22), we have:

$$\begin{aligned}
\Pr\{E(\tau, c) \mid ABC(\tau, c)\} &= 1 - \left(1 - \frac{S-2}{\Delta_{est}}\right)^{N-2} \\
&\quad \left\{ \begin{array}{l} \text{substitute } x = \frac{\Delta_{est}}{S-2} \\ p = \frac{S}{\Delta_{est}} \implies \frac{S}{\Delta_{est}} \leq \frac{1}{2} \implies \frac{\Delta_{est}}{S} \geq 2 \implies x \geq 2 \end{array} \right\} \\
&\geq 1 - \left(1 - \frac{1}{x}\right)^{N-2} \\
&\quad \left\{ \text{using (19), } N-2 \geq \frac{\Delta_{est}}{2} \implies N-2 \geq \frac{x(S-2)}{2} \right\} \\
&\geq 1 - \left\{ \left(1 - \frac{1}{x}\right)^x \right\}^{\frac{S-2}{2}} \\
&\quad \left\{ \left(1 - \frac{1}{x}\right)^x \leq \frac{1}{e} \text{ when } x \geq 2 \right\} \\
&\geq 1 - \left(\frac{1}{e}\right)^{\frac{S-2}{2}} \tag{23}
\end{aligned}$$

Using (17) and (23), we have:

$$\begin{aligned}
\Pr\{D(\tau, c) \mid ABC(\tau, c)\} &\geq \frac{1}{2} \cdot \left(\frac{S-1}{S}\right) \cdot \left\{1 - \left(\frac{1}{e}\right)^{\frac{S-2}{2}}\right\} \\
&\geq \frac{1}{2} \cdot \left(1 - \frac{1}{S}\right) \cdot \left(1 - \frac{1}{\sqrt{e}}\right)
\end{aligned}$$

$$\geq \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{1}{9}$$

Let $\sigma(N, S)$ denote the slow-down factor experienced by a neighbor discovery algorithm that is undergoing a jamming attack. The above analysis implies that $\sigma(N, S) \leq 18$ for all $N \geq 3$ and $S \geq 3$. For sufficiently large values of N and S , we show that the slow-down factor is much smaller. Again, we analyze the cases corresponding to when $p = \frac{1}{2}$ and when $p = \frac{S}{\Delta_{est}}$ separately.

- **Case 1** ($p = \frac{1}{2}$): Using (17) and (22) we have:

$$\begin{aligned} \Pr\{D(\tau, c) \mid ABC(\tau, c)\} &\geq \frac{1}{2} \cdot \left(\frac{S-1}{S}\right) \cdot \left\{1 - \left(1 - \frac{S-2}{2S}\right)^{N-2}\right\} \\ &\quad \left\{\text{for sufficiently large } S, \frac{S-1}{S} \approx 1 \text{ and } \frac{S-2}{2S} \approx \frac{1}{2}\right\} \\ &\approx \frac{1}{2} \cdot 1 \cdot \left\{1 - \left(\frac{1}{2}\right)^{N-2}\right\} \\ &\quad \left\{\text{for sufficiently large } N, \left(\frac{1}{2}\right)^{N-2} \approx 0\right\} \\ &\approx \frac{1}{2} \cdot (1 - 0) = \frac{1}{2} \end{aligned}$$

- **Case 2** ($p = \frac{S}{\Delta_{est}}$): Using (17) and (23), we have:

$$\begin{aligned} \Pr\{D(\tau, c) \mid ABC(\tau, c)\} &\geq \frac{1}{2} \cdot \left(\frac{S-1}{S}\right) \cdot \left\{1 - \left(\frac{1}{e}\right)^{\frac{S-2}{2}}\right\} \\ &\quad \left\{\text{for sufficiently large } S, \frac{S-1}{S} \approx 1 \text{ and } \left(\frac{1}{e}\right)^{\frac{S-2}{2}} \approx 0\right\} \\ &\approx \frac{1}{2} \cdot 1 \cdot (1 - 0) = \frac{1}{2} \end{aligned}$$

The above analysis implies that:

$$\lim_{\substack{N \rightarrow \infty \\ S \rightarrow \infty}} \sigma(N, S) \leq 2$$

Finally, we have:

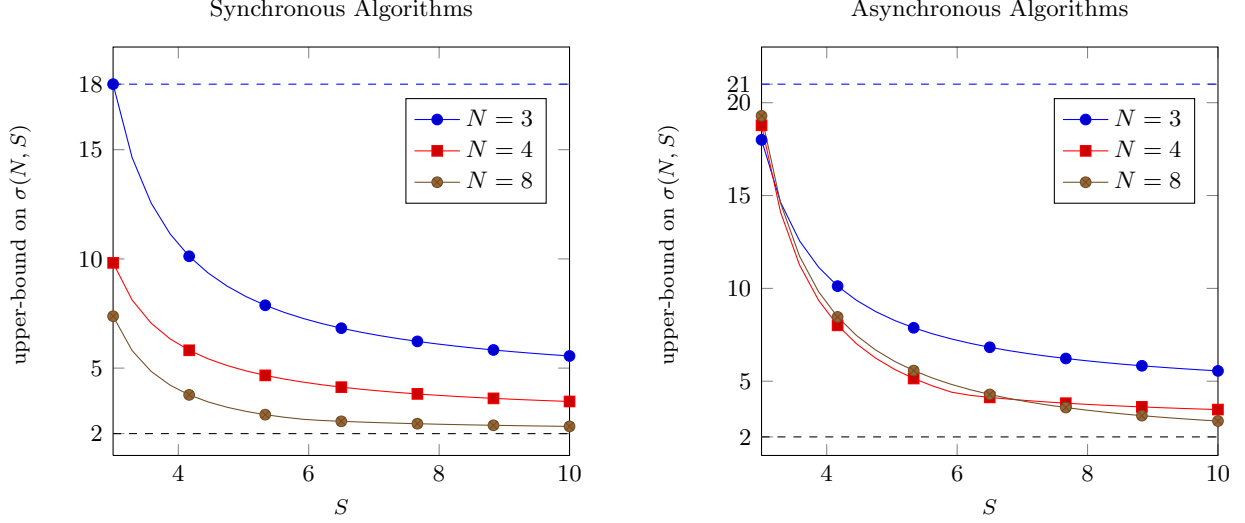


Figure 5: Upper bound on slow-down factor as a function of N and S .

Theorem 14. *The running time of the synchronous algorithms increases by a factor of 18 in the worst case. Furthermore, when N and S are sufficient large, the running time only increases by a factor of at most 2 in the worst case.*

Note that the slow-down factor of 18 represents the worst-case and only occurs for a “small” system which is easier to jam. As N and S , increases, the slow-down factor decreases rapidly as depicted in Fig. 5.

6.3.2 Asynchronous Algorithms

In the asynchronous case, we require a minor change to Algorithm 5 and Algorithm 6. Namely, when a node is in transmit mode, it makes a uniform selection of channel at the beginning of each slot, rather than at the beginning of each frame. We also assume that the length of a jamming round is at least the maximum length of a slot (measured in real-time) (which is given by $\frac{L}{3(1-\delta)}$).

Consider the link from node v to node u . Consider a legal pair of frames $\langle f, g \rangle$ with $node(f) = v$ and $node(g) = u$. Let $slot(f, g)$ denote the slot of f that lies completely within g . (If there are multiple such slots, then choose one arbitrarily.) For analyzing the jamming attack, we narrow the period during which no transmission by other neighbors of u interferes with v ’s transmission during $slot(f, g)$. This, in conjunction with the assumption made on the minimum length of jamming round, ensures that the set of channels a jammer observes to be in use in two different scans do not

have any correlation. Intuitively, this is because two consecutive scans by a jammer will involve two different slots of any node and, in transmit mode, a node now chooses the channel on which to transmit at the beginning of a slot (rather than a frame).

Consider events $\hat{A}(f, g, c)$, $\hat{B}(f, g, c)$ and $\hat{C}(f, g, c)$. To analyze our algorithms under jamming attack, we replace the event $\hat{A}(f, g, c)$ with a new event $\hat{A}'(f, g, c)$ which denotes the condition that u transmits on c during $slot(f, g)$. Further, we replace the event $\hat{C}(f, g, c)$ with a new event $\hat{C}'(f, g, c)$ which denotes the condition that no neighbor of u besides v transmits on c during any of its slots that overlaps with $slot(f, g)$. Clearly, in the absence of jamming attack, u will discover v during frame g if events $\hat{A}'(f, g, c)$, $\hat{B}(f, g, c)$ and $\hat{C}'(f, g, c)$ occur. Analogous to Lemma 5, it can be easily shown that a single slot of a node can overlap with at most three slots of another node. Note that these three slots can come from at most two different frames. It can be verified that (10) holds for $\hat{A}'(f, g, c)$ and (12) holds for $\hat{C}'(f, g, c)$ (for the latter, see Section A.7 in the appendix).

Analogous to events $ABC(\tau, c)$, $D(\tau, c)$ and $E(\tau, c)$ and $F(\tau, c, w)$ define in the synchronous case, we can define events $\hat{A}\hat{B}\hat{C}(f, g, c)$, $\hat{D}(f, g, c)$, $\hat{E}(f, g, c)$ and $\hat{F}(f, g, c, w)$, respectively. Most of the analysis for the synchronous case is also applicable to the asynchronous case. Specifically, when $p = \frac{1}{2}$, the analysis and results for the synchronous case also apply to the asynchronous case. However, when $p = \frac{S}{3\Delta_{est}}$, the results for the asynchronous case are slightly different from that for the synchronous case. Since the analysis is very similar to that for the synchronous case, we only present the main steps. Using an expression similar to (22), when $p = \frac{S}{3\Delta_{est}}$, we can show the following:

$$\begin{aligned} \Pr\{\hat{E}(f, g, c) \mid \hat{A}\hat{B}\hat{C}(f, g, c)\} &= 1 - \left(1 - \frac{S-2}{3\Delta_{est}}\right)^{N-2} \\ &\geq 1 - \left(\frac{1}{e}\right)^{\frac{S-2}{6}} \end{aligned}$$

For arbitrary N and S with $N, S \geq 3$, we have:

$$\begin{aligned} \Pr\{\hat{D}(f, g, c) \mid \hat{A}\hat{B}\hat{C}(f, g, c)\} &\geq \frac{1}{2} \cdot \left(\frac{S-1}{S}\right) \cdot \left\{1 - \left(\frac{1}{e}\right)^{\frac{S-2}{6}}\right\} \\ &\geq \frac{1}{2} \cdot \left(1 - \frac{1}{S}\right) \cdot \left(1 - \frac{1}{\sqrt[6]{e}}\right) \\ &= \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{1}{7} = \frac{1}{21} \end{aligned}$$

For sufficiently large N and S , we have:

$$\begin{aligned}\Pr\{\hat{D}(f, g, c) \mid A\hat{B}C(f, g, c)\} &\geq \frac{1}{2} \cdot \left(\frac{S-1}{S}\right) \cdot \left\{1 - \left(\frac{1}{e}\right)^{\frac{S-2}{6}}\right\} \\ &\approx \frac{1}{2} \cdot 1 \cdot (1 - 0) = \frac{1}{2}\end{aligned}$$

The above analysis implies that:

$$\forall N, \forall S : N \geq 3, S \geq 3 : \sigma(N, S) \leq 21 \quad \text{and} \quad \lim_{\substack{N \rightarrow \infty \\ S \rightarrow \infty}} \sigma(N, S) \leq 2$$

Finally, we have the following result:

Theorem 15. *The running time of the asynchronous algorithms increases by a factor of 21 in the worst case. Furthermore, when N and S are sufficient large, the running time only increases by a factor of at most 2 in the worst case.*

Again, as in the synchronous case, the slow-down factor of 21 represents the worst-case and only occurs for a “small” system which is easier to jam. As N and S , increases, the slow-down factor decreases rapidly as depicted in Fig. 5.

6.3.3 Discussion

It is possible for a jammer to jam a band of channels and not just a single channel. In that case, we require the the available channel set to contain at least three *orthogonal bands* of channels. The results derived above still hold if the analysis is conducted in terms of bands rather than channels and all bands are of the same size.

It is also possible to analyze the running time of our algorithms when a jammer can jam up to k channels in a jamming round for a fixed k , where $1 \leq k < \frac{S}{2}$. The analysis is more involved and requires stronger assumptions on system size. The slow-down factor also increases to $O(k)$ and is no longer a constant.

7 Related Work

A large number of neighbor discovery algorithms have been proposed in the literature. To our knowledge, most of the algorithms suffer from one or more of the following limitations:

- *Single-Channel Network:* Several neighbor discovery algorithms have been proposed for a (traditional) single channel wireless network (*e.g.*, [3, 4, 9, 24, 25, 31, 42, 51, 56–58, 61, 64, 67, 69]). Some of these neighbor discovery algorithms, such as those proposed in [56, 58], can be extended to work for a multi-channel network (including a heterogeneous network). The main idea is as follows. Let the collective set of all channels over which radio nodes in the network are capable of operating be referred to as the *universal channel set*. The main idea is to execute a separate instance of single-channel neighbor discovery algorithm on all channels in the universal channel set *concurrently* by interleaving their executions. A node only participates in instances that are associated with channels in its available channel set. However, this simple approach has several disadvantages. First, it requires that all nodes have to agree on the composition of the universal channel set. Second, the time complexity of the algorithm for multi-channel network (obtained as above) will always be *linear* in the size of the universal channel set. This is true even if the available channel set of all nodes contain a single common channel (an extreme case). In many cases, the available channel sets of nodes may be much smaller than the universal channel set. Third, all nodes should start executing the algorithm at the same time. Otherwise, different nodes may tune to different channels in the same time slot, thereby causing the multi-channel neighbor discovery algorithm to fail. As such, this transformation *does not work for an asynchronous system*. In comparison, in our algorithms, nodes do not need to agree on a universal channel set and the time complexity of our algorithms depends on the “degree of heterogeneity” in the network.
- *Single-Hop Network:* Neighbor discovery algorithms in [2, 5, 6] consider a single-hop network in which all nodes can communicate with each other directly. Our algorithms, on the other hand, work even for a multi-hop network.
- *Synchronous System:* Neighbor discovery algorithms in [17, 33, 55] assume a time-slotted synchronous system. Work in [35, 44, 66] assume a synchronous system in which all nodes initiate

neighbor discovery at the same time. Moreover, the proposed algorithms are deterministic in nature, and have high time complexity that depends on the *product* of network size and universal channel set size. In contrast, our focus in this work is on developing algorithms for an asynchronous system that have lower running time (albeit provide only probabilistic guarantees).

- *Homogeneous Channel Availability:* Neighbor discovery algorithms in [2, 17, 34] assume homogeneous channel availability (all channels are available to all nodes). Our algorithms, on the other hand, work even when different nodes have different subsets of channels available for communication, and, further, their running time adapts to the “degree of heterogeneity” in the network.
- *Little or No Theoretical Guarantees:* Work in [48, 60] only evaluate the proposed neighbor discovery algorithms using experiments; no attempt is made to provide any theoretical (deterministic or probabilistic) guarantees. Work in [6] provides a suite of neighbor discovery algorithms that can tolerate jamming attacks. However, it does not provide any theoretical guarantees either (algorithms are evaluated using only simulation experiments). In our work, in contrast, we analyze all our algorithms theoretically and establish probabilistic bounds on their performance.

Besides the related work mentioned above, Raniwala and Chiueh propose a neighbor discovery algorithm for a multi-channel wireless network in [49] but their work assumes that each node has multiple interfaces. Law *et al.* describe a neighbor discovery algorithm for constructing a scatternet in [36]. Salondis *et al.* describe an algorithm for *two-node* link formation in [52], which is then used as a building block for forming a scatternet in [53]. A scatternet is an ad hoc network consisting of two or more piconets; a piconet is an ad hoc network consisting two or more (up to eight) Bluetooth-enabled devices with one device acting as master and remaining devices acting as slaves. Work in [8, 17, 19, 20, 28, 30, 41, 43] addresses the closely-related problem of *rendezvous* between two CR nodes. Intuitively, the rendezvous problem between two nodes involves ensuring that the two nodes tune to the same channel at the same time, which is necessary for nodes to discover each other. However, as opposed to the problem considered in this work, the rendezvous problem does

not consider the effect of possible simultaneous transmissions by other neighboring nodes at the same, which may cause collisions thereby preventing a link from being discovered.

8 Conclusion

In this work, we have proposed several randomized neighbor discovery algorithms for an M²HeW network both for synchronous and asynchronous systems under a variety of assumptions. Our algorithms guarantee success with arbitrarily high probability. All our algorithms are designed to minimize the amount of knowledge that nodes need to possess *a priori* before they start executing the neighbor discovery algorithm. We have shown that our neighbor discovery algorithms are robust to unreliable channels are adversarial attacks.

As a future, we plan to develop randomized algorithms for solving other communication problems for an M²HeW network such as broadcasting and convergecasting that guarantee success with arbitrarily high probability.

References

- [1] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan. Cooperative Spectrum Sensing in Cognitive Radio Networks: A Survey. *Physical Communication*, 4(1):40–62, March 2011.
- [2] G. Alonso, E. Kranakis, C. Sawchuk, R. Wattenhofer, and P. Widmayer. Probabilistic Protocols for Node Discovery in Ad-Hoc Multi-Channel Broadcast Networks. In *Proceedings of the 2nd Annual Conference on Adhoc Networks and Wireless (ADHOCNOW)*, pages 101–115, 2003.
- [3] G. Alonso, E. Kranakis, R. Wattenhofer, and P. Widmayer. Probabilistic Protocols for Node Discovery in Ad-Hoc, Single Broadcast Channel Networks. In *Proceedings of the International Parallel and Distributed Processing Symposium (IPDPS)*, Nice, France, April 2003.
- [4] X. An and R. Hekmat. Self-Adaptive Neighbor Discovery in Ad Hoc Networks with Directional Antennas. In *Proceedings of the 16th IST Summit on Mobile and Wireless Communications*, pages 1–5, 2007.

- [5] C. Arachchige, S. Venkatesan, and N. Mittal. An Asynchronous Neighbor Discovery Algorithm for Cognitive Radio Networks. In *Proceedings of the 3rd IEEE Symposia on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Chicago, Illinois, USA, 2008.
- [6] A. Asterjadhi and M. Zorzi. JENNA: A Jamming Evasive Network-coding Neighbor-discovery Algorithm for Cognitive Radio Networks. In *Proceedings of the 2010 IEEE International Conference on Communications Workshops*, Capetown, South Africa, May 2010.
- [7] L. Bao and J. J. Garcia-Luna-Aceves. Transmission Scheduling in Ad Hoc Networks with Directional Antennas. In *Proceedings of the 8th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, pages 48–58, Atlanta, Georgia, USA, September 2002.
- [8] K. Bian and J.-M. Park. Maximizing Rendezvous Diversity in Rendezvous Protocols for Decentralized Cognitive Radio Networks. *IEEE Transactions on Mobile Computing (TMC)*, 12(7):1294–1307, July 2013.
- [9] S. A. Borbash, A. Ephremides, and M. J. McGlynn. An Asynchronous Neighbor Discovery Algorithm for Wireless Sensor Networks. *Ad Hoc Networks*, 5(7):998–1016, September 2007.
- [10] A. Boukerche. *Handbook of Algorithms For Wireless Networking and Mobile Computing*. CRC Press, 2005.
- [11] A. Boukerche. *Algorithms and Protocols for Wireless, Mobile Ad Hoc Networks*, volume 77. John Wiley and Sons, Incorporated, 2008.
- [12] A. Boukerche, C. Rezende, and R. W. Pazzi. Improving Neighbor Localization in Vehicular Ad Hoc Networks to Avoid Overhead from Periodic Messages. In *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, November 2009.
- [13] R. W. Broderson, A. Wolisz, D. Cabric, S. M. Mishra, and D. Willkomm. CORVUS: A Cognitive Radio Approach for Usage of Virtual Unlicensed Spectrum. Available at http://bwrc.eecs.berkeley.edu/Research/MCMA/CR_White_paper_final1.pdf.

- [14] G. Cafaro, N. Correal, and D. Taubenheim. A 100 MHz – 2.5 GHz CMOS Transceiver in an Experimental Cognitive Radio System. In *Proceedings of the SDR (Software Defined Radio) Technical Conference and Product Exposition*, Denver, Colorado, USA, November 2007.
- [15] M. Cesana, F. Cuomo, and E. Ekici. Routing in Cognitive Radio Networks: Challenges and Solutions. *Ad Hoc Networks*, 9(3):228–248, May 2010.
- [16] Y.-L. Chang and C.-C. Hsu. Routing in Wireless/Mobile Ad-Hoc Networks via Dynamic Group Construction. *Mobile Networks and Applications (MONET)*, 5(1):27–37, March 2000.
- [17] C.-M. Chao and C.-Y. Hsu. Supporting Fast Neighbor Discovery for Cognitive Radio Networks. In *Proceedings of the International Conference on Computational Problem-Solving (ICCP)*, pages 123–126, October 2013.
- [18] C. Chen, M. Song, C.S. Xin, and J. Backens. A Game-Theoretical Anti-Jamming Scheme for Cognitive Radio Networks. *IEEE Network*, 27(3):22–27, May/June 2013.
- [19] L. Chen, K. Bian, L. Chen, C. Liu, J.-M. J. Park, and X. Li. A Group-Theoretic Framework for Rendezvous in Heterogeneous Cognitive Radio Networks. In *Proceedings of the 15th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 165–174, August 2014.
- [20] L. Chen, K. Bian, and M. Zheng. Heterogeneous Multi-Channel Neighbor Discovery for Mobile Sensing Applications: Theoretical Foundation and Protocol Design. In *Proceedings of the 15th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 307–316, August 2014.
- [21] Y. Choi, M. Khan, V. S. Anil Kumar, and G. Pandurangan. Energy-Optimal Distributed Algorithms for Minimum Spanning Trees. *IEEE Journal on Selected Areas in Communications (JSAC)*, 27(7):1297–1304, September 2009.
- [22] R. Choudhury, X. Yang, R. Ramanathan, and N. Vaidya. Using Directional Antennas for Medium Access Control in Ad Hoc Networks. In *Proceedings of the 8th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, pages 59–70, Atlanta, Georgia, USA, September 2002.

- [23] K. Dabcevic, A. Betancourt, L. Marcenaro, and C. S. Regazzoni. A Fictitious Play-Based Game-Theoretical Approach to Alleviating Jamming Attacks for Cognitive Radios. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8158–8162, Florence, Italy, May 2014.
- [24] P. Dutta and D. Culler. Practical Asynchronous Neighbor Discovery and Rendezvous for Mobile Sensing Applications. In *Proceedings of the ACM Conference on Embedded Networked Sensor Systems (SenSys)*, pages 71–84, November 2008.
- [25] V. Dyo and C. Mascolo. Efficient Node Discovery in Mobile Wireless Sensor Networks. In *Proceedings of the IEEE international Conference on Distributed Computing in Sensor Systems (DCOSS)*, pages 478–485, Santorini Island, Greece, June 2008.
- [26] S. Gandham, M. Dawande, and R. Prakash. Link Scheduling in Wireless Sensor Networks: Distributed Edge-Coloring Revisited. *Journal of Parallel and Distributed Computing (JPDC)*, 68(8):1122–1134, August 2008.
- [27] S. Gandham, Y. Zhang, and Q. Huang. Distributed Time-Optimal Scheduling for Convergecast in Wireless Sensor Networks. *Computer Networks (COMNET)*, 52:610–629, 2008.
- [28] R. Gandhi, C.-C. Wang, and Y. C. Hu. Fast Rendezvous for Multiple Clients for Cognitive Radios using Coordinated Channel Hopping. In *Proceedings of the 8th IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pages 434–442, June 2012.
- [29] A. Gonga, T. Charalambous, and M. Johansson. Neighbor Discovery in Multichannel Wireless Clique Networks: An Epidemic Approach. In *Proceedings of the 10th IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pages 131–135, 2013.
- [30] Z. Gu, Q.-S. Hua, Y. Wang, and F. C. M. Lau. Nearly Optimal Asynchronous Blind Rendezvous Algorithm for Cognitive Radio Networks. In *Proceedings of the 10th IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pages 371–379, June 2013.

- [31] E. B. Hamida, G. Chelius, A. Busson, and E. Fleury. Neighbor Discovery in Multi-Hop Wireless Networks: Evaluation and Dimensioning with Interferences Considerations. *Discrete Mathematics and Theoretical Computer Science*, 10(2):87–114, 2008.
- [32] W. Heinzelman, A. Chandrasekaran, and H. Balakrishnan. Energy Efficient Communication Protocol for Wireless Microsensor Networks. In *Proceedings of the Hawaii International Conference on Systems Sciences*, pages 3005–3014, Maui, Hawaii, USA, January 2000.
- [33] V. Iyer, A. Pruteanu, and S. Dulman. NetDetect: Neighborhood Discovery in Wireless Networks Using Adaptive Beacons. In *Proceedings of the 5th IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*, pages 31–40, October 2011.
- [34] N. Karowski, A. C. Viana, and A. Wolisz. Optimized Asynchronous Multi-Channel Neighbor Discovery. In *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, pages 536–540, 2011.
- [35] S. Krishnamurthy, M. Thoppian, S. Kuppa, R. Chandrasekaran, N. Mittal, S. Venkatesan, and R. Prakash. Time-efficient Distributed Layer-2 Auto-configuration for Cognitive Radio Networks. *Computer Networks (COMNET)*, 52(4):831–849, March 2008.
- [36] C. Law, A. K. Mehta, and K.-Y. Siu. Performance of a new Bluetooth Scatternet Formation Protocol. In *Proceedings of the 2nd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 183–192, 2001.
- [37] L. Li, J. Y. Halpern, P. Bahl, Y.-M. Wang, and R. Wattenhofer. A Cone-Based Distributed Topology-Control Algorithm for Wireless Multi-Hop Networks. *IEEE/ACM Transactions on Networking*, 13(1):147–159, 2005.
- [38] X. Li and W. Cadeau. Anti-Jamming Performance of Cognitive Radio Networks. In *Proceedings of the 45th Annual Conference on Information Systems and Sciences (CISS)*, Baltimore, Maryland, USA, March 2011.
- [39] H. Lim and C. Kim. Multicast Tree Construction and Flooding in Wireless Ad Hoc Networks. In *Proceedings of the 3rd ACM International Workshop on Modeling, Analysis and Simulation*

- of *Wireless and Mobile Systems (MSWiM)*, pages 61–68, Boston, Massachusetts, USA, August 2000.
- [40] C. R. Lin and M. Gerla. Adaptive Clustering for Mobile Wireless Networks. *IEEE Journal on Selected Areas in Communications (JSAC)*, 15(7):1265–1275, September 1997.
 - [41] H. Liu, Z. Lin, X. Chu, and Y.-W. Leung. Jump-Stay Rendezvous Algorithm for Cognitive Radio Networks. *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, 23(10):1867–1881, October 2012.
 - [42] M. J. McGlynn and S. A. Borbash. Birthday Protocols for Low Energy Deployment and Flexible Neighbor Discovery in Ad Hoc Wireless Networks. In *Proceedings of the 2nd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 137–145, Long Beach, California, USA, 2001.
 - [43] J. Misić, V. B. Misić, M. S. I. Khan, and M. M. Rahman. Properties of Blind Rendezvous in Channel Hopping Cognitive Piconets. In *Proceedings of the IEEE 78th Vehicular Technology Conference (VTC Fall)*, September 2013.
 - [44] N. Mittal, S. Krishnamurthy, R. Chandrasekaran, S. Venkatesan, and Y. Zeng. On Neighbor Discovery in Multi-Channel Cognitive Radio Networks. *Journal of Parallel and Distributed Computing (JPDC)*, 69(7):623–637, July 2009.
 - [45] N. Mittal, Y. Zeng, S. Venkatesan, and R. Chandrasekaran. Randomized Distributed Algorithms for Neighbor Discovery in Multi-hop Multi-channel Heterogeneous Wireless Networks. In *Proceedings of the 31st IEEE International Conference on Distributed Computing Systems (ICDCS)*, pages 57–66, Minneapolis, Minnesota, USA, June 2011.
 - [46] K. Nakano and S. Olariu. Randomized Leader Election Protocols in Radio Networks with no Collision Detection. In *Proceedings of the Annual International Symposium on Algorithms and Computation (ISAAC)*, pages 362–373, Taipei, Taiwan, December 2000.
 - [47] R. Prakash. Unidirectional Links Prove Costly in Wireless Ad Hoc Networks. In *Proceedings of the 3rd International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications (DIAL-M)*, pages 15–22, Seattle, Washington, USA, August 1999.

- [48] L. Raju, S. Ganu, B. Anepu, I. Seskar, and D. Raychaudhuri. Beacon Assisted Discovery Protocol (BEAD) for Self-Organizing Hierarchical Ad-Hoc Networks. In *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, volume 3, pages 1676–1680, 2004.
- [49] A. Raniwala and T. c. Chiueh. Architecture and Algorithms for an IEEE 802.11-based Multi-Channel Wireless Mesh Network. In *Proceedings of the 24th IEEE Conference on Computer Communications (INFOCOM)*, pages 2223–2234, Miami, Florida, USA, March 2005.
- [50] M. H. Rehmania, A. C. Vianab, H. Khalifec, and S. Fdidaa. SURF: A Distributed Channel Selection Strategy for Data Dissemination in Multi-Hop Cognitive Radio Networks. *Computer Communications (COMCOM)*, 36(10-11):1172–1185, June 2013.
- [51] A. Russell, S. Vasudevan, B. Wang, W. Zeng, X. Chen, and W. Wei. Neighbor Discovery in Wireless Networks with Multipacket Reception. *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, 2014. PrePrint, doi:10.1109/TPDS.2014.2321157.
- [52] T. Salonidis, P. Bhagwat, and L. Tassiulas. Proximity Awareness and Fast Connection Establishment in Bluetooth. In *Proceedings of the 1st ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 141–142, 2000.
- [53] T. Salonidis, P. Bhagwat, L. Tassiulas, and R. LaMaire. Distributed Topology Construction of Bluetooth Personal Area Networks. In *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, pages 1577–1586, April 2001.
- [54] Y. Song and J. Xie. A Distributed Broadcast Protocol in Multi-Hop Cognitive Radio Ad Hoc Networks without a Common Control Channel. In *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, pages 2273–2281, March 2012.
- [55] G. Sun, F. Wu, X. Gao, and G. Chen. PHED: Pre-Handshaking Neighbor Discovery Protocols in full duplex wireless ad hoc networks. In *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, pages 584–590, Dec 2012.

- [56] S. Vasudevan, M. Adler, D. Goeckel, and D. Towsley. Efficient Algorithms for Neighbor Discovery in Wireless Networks. *IEEE/ACM Transactions on Networking*, 21(1):69–83, February 2013.
- [57] S. Vasudevan, J. Kurose, and D. Towsley. On Neighbor Discovery in Wireless Networks with Directional Antennas. In *Proceedings of the 24th IEEE Conference on Computer Communications (INFOCOM)*, pages 2502–2512, Miami, Florida, USA, March 2005.
- [58] S. Vasudevan, D. Towsley, D. Goeckel, and R. Khalili. Neighbor Discovery in Wireless Networks and the Coupon Collector’s Problem. In *Proceedings of the 15th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, pages 181–192, Beijing, China, 2009.
- [59] B. Wang, Y. Wu, K. J. R. Liu, and T. C Clancy. An Anti-Jamming Stochastic Game for Cognitive Radio Networks. *IEEE Journal on Selected Areas in Communications (JSAC)*, 29(4):877–889, April 2011.
- [60] F. Wang, S. Sun, L. Li, W. Guo, and Y. Ma. Neighbor Discovery based on Group Frequency Hopping without Common Control Channel for Cognitive Radio Ad Hoc Networks. In *Proceedings of the 15th IEEE International Conference on Communication Technology (ICCT)*, pages 483–487, Guilin, China, 2013.
- [61] K. Wang, X. Mao, and Y. Liu. BlindDate: A Neighbor Discovery Protocol. *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, 26(4):949–959, April 2015.
- [62] R. Wattenhofer, L. Li, P. Bahl, and Y.-M. Wang. Distributed Topology Control for Power Efficient Operation in Multihop Wireless Ad Hoc Networks. In *Proceedings of the 20th IEEE Conference on Computer Communications (INFOCOM)*, pages 1388–1397, Anchorage, Alaska, USA, April 2001.
- [63] A. Willig, N. Karowski, and J.-H. Hauer. Passive Discovery of IEEE 802.15.4-based Body Sensor Networks. *Annual Conference on Adhoc Networks and Wireless (ADHOCNOW)*, 8(7):742–754, September 2010.

- [64] D. Yang, J. Shin, J. Kim, and C. Kim. Asynchronous Probing Scheme for the Optimal Energy-Efficient Neighbor Discovery in Opportunistic Networking. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 1–4, Galveston, Texas, USA, March 2009.
- [65] T. Yücek and H. Arslan. A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications. *IEEE Communications Surveys & Tutorials*, 11(1):116–130, 2009.
- [66] Y. Zeng, N. Mittal, S. Venkatesan, and R. Chandrasekaran. Fast Neighbor Discovery with Lightweight Termination Detection in Heterogeneous Cognitive Radio Networks. In *Proceedings of the 9th International Symposium on Parallel and Distributed Computing (ISPDC)*, pages 149–156, Istanbul, Turkey, July 2010.
- [67] L. Zhang, J. Luo, and D. Guo. Neighbor Discovery for Wireless Networks via Compressed Sensing. *Performance Evaluation*, 70(7):457–471, 2013.
- [68] R. Zhang and Y. Zhang. Wormhole-Resilient Secure Neighbor Discovery in Underwater Acoustic Networks. In *Proceedings of the 29th IEEE Conference on Computer Communications (INFOCOM)*, pages 1–9, March 2010.
- [69] R. Zheng, J. Hou, and L. Sha. Asynchronous Wakeup for Ad Hoc Networks. In *Proceedings of the 4th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 35–45, Annapolis, Maryland, USA, June 2003.

A Omitted Proofs

A.1 Computing the probability of occurrence of $A(\tau, c)$

We have:

$$\begin{aligned}
 \Pr\{A(\tau, c)\} &= (v \text{ selects } c \text{ at the beginning of } \tau) \wedge (v \text{ chooses to transmit during } \tau) \\
 &= \frac{1}{|\mathcal{A}(v)|} \times \min\left(\frac{1}{2}, \frac{|\mathcal{A}(v)|}{2^k}\right) \\
 &= \min\left(\frac{1}{2|\mathcal{A}(v)|}, \frac{1}{2^k}\right)
 \end{aligned}$$

$$\begin{aligned}
& \left\{ \text{using (2), } 2^{k-1} \leq \Delta(u, c) \text{ which implies } 2^k \leq 2\Delta(u, c) \right\} \\
& \geq \frac{1}{\max\{2|\mathcal{A}(v)|, 2\Delta(u, c)\}} \\
& \geq \frac{1}{2\max(S, \Delta)}
\end{aligned}$$

A.2 Computing the probability of occurrence of $B(\tau, c)$

We have:

$$\begin{aligned}
\Pr\{B(\tau, c)\} &= (u \text{ selects } c \text{ at the beginning of } \tau) \wedge (u \text{ chooses to listen during } \tau) \\
&= \frac{1}{|\mathcal{A}(u)|} \times \left\{ 1 - \min\left(\frac{1}{2}, \frac{|\mathcal{A}(u)|}{2^k}\right) \right\} \\
&\quad \left\{ \min(x, y) \leq x \right\} \\
&\geq \frac{1}{|\mathcal{A}(u)|} \times \left(1 - \frac{1}{2}\right) = \frac{1}{2|\mathcal{A}(u)|}
\end{aligned}$$

A.3 Computing the probability of occurrence of $C(\tau, c)$

Let $\mathcal{N}(u, c)$ denote the set of neighbors of u on c . Note that, if $\mathcal{N}(u, c)$ only contains v , then $\Pr\{C(\tau, c)\} = 1$. Otherwise, we have:

$$\begin{aligned}
\Pr\{C(\tau, c)\} &= \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \Pr(w \text{ does not transmit on } c \text{ during } \tau) \\
&= \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \{1 - \Pr(w \text{ transmits on } c \text{ during } \tau)\} \\
&= \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \left\{ 1 - \frac{1}{|\mathcal{A}(w)|} \times \min\left(\frac{1}{2}, \frac{|\mathcal{A}(w)|}{2^k}\right) \right\} \\
&= \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \left\{ 1 - \min\left(\frac{1}{2|\mathcal{A}(w)|}, \frac{1}{2^k}\right) \right\} \\
&\quad \left\{ \min(x, y) \leq y \right\} \\
&\geq \prod_{\substack{w \in \mathcal{N}(u, c) \\ w \neq v}} \left(1 - \frac{1}{2^k}\right)
\end{aligned}$$

$$\begin{aligned}
&= \left(1 - \frac{1}{2^k}\right)^{|\mathcal{N}(u,c)|-1} \\
&\quad \left\{ |\mathcal{N}(u,c)| - 1 = \Delta(u,c) - 1 \text{ and, using (2), } \Delta(u,c) - 1 \leq 2^k \right\} \\
&\geq \left(1 - \frac{1}{2^k}\right)^{2^k} \\
&\quad \left\{ \forall x \geq 2, \left(1 - \frac{1}{x}\right)^x \text{ is a monotonically increasing function of } x \text{ and thus } \geq \frac{1}{4} \right\} \\
&\geq \frac{1}{4}
\end{aligned}$$

A.4 Computing the probability that s covers (v, u)

Let $F(\tau, c)$ denote the event that τ covers (v, u) on c . Note that $F(\tau, c) = A(\tau, c) \wedge B(\tau, c) \wedge C(\tau, c)$. Since $A(\tau, c)$, $B(\tau, c)$ and $C(\tau, c)$ are mutually independent events, $\Pr\{F(\tau, c)\} = \Pr\{A(\tau, c)\} \times \Pr\{B(\tau, c)\} \times \Pr\{C(\tau, c)\}$. We have:

$$\begin{aligned}
\Pr\{F(\tau, c)\} &\geq \frac{1}{2 \max(S, \Delta)} \times \frac{1}{2|\mathcal{A}(u)|} \times \frac{1}{4} \\
&= \frac{1}{16|\mathcal{A}(u)| \max(S, \Delta)}
\end{aligned}$$

Now, we have:

$$\begin{aligned}
&\Pr\{s \text{ covers } (v, u)\} \\
&= \sum_c \Pr\{s \text{ covers } (v, u) \text{ on } c\} \\
&= \sum_{c \in \mathcal{A}(v) \cap \mathcal{A}(u)} \Pr\{s \text{ covers } (v, u) \text{ on } c\} + \sum_{c \notin \mathcal{A}(v) \cap \mathcal{A}(u)} \Pr\{s \text{ covers } (v, u) \text{ on } c\} \\
&\quad \left\{ c \notin \mathcal{A}(v) \cap \mathcal{A}(u) \text{ implies that } \Pr\{s \text{ covers } (v, u) \text{ on } c\} = 0 \right\} \\
&= \sum_{c \in \mathcal{A}(v) \cap \mathcal{A}(u)} \Pr\{s \text{ covers } (v, u) \text{ on } c\} \\
&\geq \sum_{c \in \mathcal{A}(v) \cap \mathcal{A}(u)} \Pr\{F(\tau, c)\} \\
&\geq \sum_{c \in \mathcal{A}(v) \cap \mathcal{A}(u)} \frac{1}{16|\mathcal{A}(u)| \max(S, \Delta)} \\
&= \frac{|\mathcal{A}(v) \cap \mathcal{A}(u)|}{|\mathcal{A}(u)|} \times \frac{1}{16 \max(S, \Delta)} \geq \frac{\rho}{16 \max(S, \Delta)}
\end{aligned}$$

A.5 Computing the probability of successful discovery of a link in $\frac{16 \max(S, \Delta)}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$ stages

Let $M = \frac{16 \max(S, \Delta)}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)$. We have:

$$\begin{aligned}
& \Pr((v, u) \text{ is not covered within } M \text{ stages}) \\
&= \prod_{1 \leq i \leq M} \Pr((v, u) \text{ is not covered in the } i^{th} \text{ stage}) \\
&\leq \prod_{1 \leq i \leq M} \left(1 - \frac{\rho}{16 \max(S, \Delta)} \right) \\
&= \left(1 - \frac{\rho}{16 \max(S, \Delta)} \right)^{\frac{16 \max(S, \Delta)}{\rho} \ln \left(\frac{N^2}{\epsilon} \right)} \\
&\quad \{\forall x \geq 2, (1 - \frac{1}{x})^x \text{ is upper bounded by } \frac{1}{e}\} \\
&\leq \left(\frac{1}{e} \right)^{\ln \left(\frac{N^2}{\epsilon} \right)} = \frac{\epsilon}{N^2}
\end{aligned}$$

A.6 Computing the sum of the number of time-slots/frames in the first k epochs

Let S denote the sum $\sum_{1 \leq i \leq k} (i+1) \cdot 2^i$. We have:

$$\begin{aligned}
S &= 2 \cdot 2^1 + 3 \cdot 2^2 + 4 \cdot 2^3 + \dots + (k+1) \cdot 2^k \\
2S &= 2 \cdot 2^2 + 3 \cdot 2^3 + \dots + k \cdot 2^k + (k+1) \cdot 2^{k+1}
\end{aligned}$$

Subtracting the first equation from the second and rearranging terms, we obtain:

$$\begin{aligned}
S &= (k+1) \cdot 2^{k+1} - (2^k + 2^{k-1} + \dots + 2^2) - 2 \cdot 2^1 \\
&= (k+1) \cdot 2^{k+1} - \left(\sum_{1 \leq i \leq k} 2^i \right) - 2 \\
&= (k+1) \cdot 2^{k+1} - 2 \cdot (2^k - 1) - 2 \\
&= k \cdot 2^{k+1}
\end{aligned}$$

A.7 Computing the probability of occurrence of $\hat{C}'(f, g, c)$

We focus on the case when exactly three slots of w overlap with $\text{slot}(f, g)$. But the result also holds if fewer than three slots of w overlap with $\text{slot}(f, g)$. There are two sub-cases: the three (overlapping) slots belong to the same frame or two different frames.

Sub-case 1 (three slots belong to the same frame): In this case, we have:

$$\begin{aligned}
 \Pr\{w \text{ does not transmit on } c \text{ during } \text{slot}(f, g)\} &= 1 - p + p \cdot \left(\frac{S-1}{S}\right)^3 \\
 &= 1 - p \left(1 - \frac{(S-1)^3}{S^3}\right) \\
 &= 1 - p \left(\frac{3S^2 - 3S + 1}{S^3}\right) \\
 &\geq 1 - p \left(\frac{3S^2}{S^3}\right) \\
 &\geq 1 - \left(\frac{S}{3\Delta_{est}}\right) \cdot \left(\frac{3}{S}\right) \\
 &= 1 - \frac{1}{\Delta_{est}}
 \end{aligned}$$

Sub-case 2 (the three slots come from two different frames): In this case, we have:

$$\begin{aligned}
 \Pr\{w \text{ does not transmit on } c \text{ during } \text{slot}(f, g)\} &= \left\{1 - p + p \cdot \left(\frac{S-1}{S}\right)\right\} \cdot \left\{1 - p + p \cdot \left(\frac{S-1}{S}\right)^2\right\} \\
 &= \left(1 - p \cdot \frac{1}{S}\right) \cdot \left(1 - p \cdot \frac{2S-1}{S^2}\right) \\
 &= 1 - p \cdot \frac{1}{S} - p \cdot \frac{2S-1}{S^2} + p^2 \cdot \frac{2S-1}{S^3} \\
 &\geq 1 - p \cdot \frac{3S-1}{S^2} \\
 &\geq 1 - p \cdot \frac{3S}{S^2} \\
 &= 1 - p \cdot \frac{3}{S} \\
 &\geq 1 - \frac{S}{3\Delta_{est}} \cdot \frac{3}{S} = 1 - \frac{1}{\Delta_{est}}
 \end{aligned}$$

It can now be shown, in a similar manner as in the proof of Lemma 6, that $\Pr\{\hat{C}'(f, g, c)\} \geq \frac{1}{4}$.